

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, 15/62, C07K 14/47, 16/18, A61K 38/17, C12N 5/10		A2	(11) International Publication Number: WO 99/32622 (43) International Publication Date: 1 July 1999 (01.07.99)
(21) International Application Number: PCT/US98/26793		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 16 December 1998 (16.12.98)		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(30) Priority Data: 08/994,494 19 December 1997 (19.12.97) US			
(71) Applicant: ZYMOGENETICS, INC. [US/US]; 1201 Eastlake Avenue East, Seattle, WA 98102 (US).			
(72) Inventor: HOLLOWAY, James, L.; 835 N.E. 89th Street, Seattle, WA 98115 (US).			
(74) Agent: LINGENFELTER, Susan, E.; ZymoGenetics, Inc., 1201 Eastlake Avenue East, Seattle, WA 98102 (US).			

(54) Title: SEMAPHORIN ZSMF-3

(57) Abstract

Semaphorin polypeptides, polynucleotides encoding the polypeptides, and related compositions and methods are disclosed. The polypeptides include a soluble, truncated semaphorin that is highly expressed in stomach tissue. The polypeptides may be used within methods for detecting receptors that mediate neurite outgrowth, and may also have immunological value.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

5
DESCRIPTION
SEMAPHORIN ZSMF-3

BACKGROUND OF THE INVENTION

Neuronal cell outgrowths, known as processes, grow away from the cell body to form synaptic connections. Long, thin processes which carry information away from the cell body are called axons, and short, thicker processes which carry information to and from the cell body are called dendrites. Axons and dendrites are collectively referred to as neurites. Neurites are extended by means of growth cones, the growing tip of the neurite, which is highly motile and is ultimately responsible for increasing and extending the neuronal network in the body. The growth cones are able to navigate their way to their targets using environmental cues or signals, which encourage or discourage the growth cone from extending the neurite in a particular direction. Such cues and signals include older neurons and orienting glial fibers, chemicals such as nerve growth factor released by astrocytes and other attracting or repelling substances released by target cells. The membrane of the growth cone bears molecules such as N-CAM (nerve cell adhesion molecule) which are attracted or repelled by environmental cues and thus influence the direction and degree of neurite growth. The growth cone also engulfs molecules from the environment which are transported to the cell body and influence growth. A number of proteins from vertebrates and invertebrates have been identified as influencing the guidance of neurite growth, either through repulsion or chemoattraction. Among those molecules are netrins, EPH-related receptor tyrosine kinases and their ligands, vitronectin, thrombospondin, human neuronal attachment factor-1 (NAF-1), connectin, adhesion molecules such as CAM (cell adhesion molecule) and

the semaphorins/collapsins (Neugebauer et al., Neuron 6:345-58, 1991; O'Shea et al., Neuron 7:231-7, 1991; Osterhout et al., Devel. Biol. 150:256-65, 1992; Goodman, Cell 78:353-6, 1993; DeFreitas et al., Neuron 15:333-43, 1995; Dodd and Schuchardy Cell 81:471-4, 1995; Keynes and Cook, Cell 83:161-9, 1995; Müller et al., Cur. Opin. Genet. and Devel. 6:469-74, 1996, Goodman, Annu. Rev. Neurosci. 19:341-77, 1996; WIPO Patent Application No: 97/29189 and Goodman et al., US Patent No. 5,639,856).

10 Semaphorins/collapsins are a family of related transmembrane and secreted molecules. Invertebrate, vertebrate and viral semaphorins are known (Kolodkin et al., Cell 75:1389-99, 1993; Luo et al., Cell 75:217-27, 1993; Ensser and Fleckenstein, J. Gen. Virol. 76:1063-7, 1995; Luo et al., Neuron 14:1131-40, 1995; Adams et al., Mech. Devel. 57:33-45, 1996; Hall et al., Proc. Natl. Acad. Sci. USA 93:11780-8, 1996; Roche et al., Oncogene 12:1289-97, 1996; Skeido et al., Proc. Natl. Acad. Sci. USA 93:4120-5, 1996; Xiang et al., Genomics 32:39-48, 1996; 15 Eckhardt et al., Mol. Cell Neurosci. 9:409-19, 1997, and Zhou et al., Mol. Cell. Neurosci. 9:26-41, 1997). Invertebrate semaphorins include but are not limited to G-sema I (grasshopper semaphorin), D-sema I and D-sema II (Drosophila semaphorin) and T-sema I (Tribolium 20 semaphorin). Vertebrate semaphorins include but are not limited to Sema A - Sema G, the collapsins 1-5, murine Sema IVa and IVb, and human sema III, human semaphorin A(V), human semaphorin IV, human semaphorin I and CD 100. Viral semaphorins include Sema IV from both vaccinia and variola 25 viruses and a semaphorin-like gene (ahv-sema) has been reported from alcelaphine herpesvirus type 1.

The semaphorins comprise a variable region of 30-60 amino acids that includes a signal sequence, followed by a conserved region of about 500 amino acid residues called 30 the semaphorin (Sema) domain. This extracellular domain contains between 14-16 conserved cysteine residues, an N-linked glycosylation site and numerous blocks of conserved

amino acids. Membrane-bound and soluble semaphorin proteins have been described, see for example, Kolodkin et al., *ibid.*; Adams et al., *ibid.* and Goodman et al., US Patent No:5,639,856. Classification of five subgroups within the semaphorin family has made based on the sequence C terminal of the semaphorin domain. Both soluble (lacking a transmembrane domain) and membrane-bound (having a transmembrane domain and localized to the membrane) semaphorins have been described. Group I includes semaphorins having a transmembrane domain followed by a cytoplasmic domain. Group II and III have no transmembrane domain, but have a region with Ig homology, Group IV has a region of Ig homology C-terminal of the semaphorin domain followed by a transmembrane and cytoplasmic domain. Group V has a series of thrombospondin repeats C-terminal to the semaphorin domain followed by a transmembrane and cytoplasmic domain.

Most insect semaphorins are membrane bound proteins and belong to Group I. G-Sema I, T-Sema I and D-Sema I have a region of 80 amino acid residues following the semaphorin domain, which is followed by a transmembrane domain and an 80-110 amino acid cytoplasmic domain. Murine Sema IVa has a transmembrane domain followed by a 216 amino acid cytoplasmic domain. Murine sema F and G are members of Group V and have an approximately 400 amino acid thrombospondin repeat domain between the semaphorin and transmembrane domains. Group II secreted proteins, such as D-sema II, have no membrane linkage but instead have a block of less than 20 amino acids following the semaphorin domain which is followed by an Ig domain and a 70 amino acid C-terminal region. Group III secreted proteins, such as H-Sema III, are similar to Group II except the C-terminal region is longer, in excess of 100 amino acids. The viral semaphorin-like gene, ahv-sema, is truncated immediately following the Ig-domain. The viral semaphorin, sema IV, has a truncated, 441 amino acid residue, semaphorin domain and no Ig region. See Kolodkin et al.,

ibid.; Adams et al. *ibid.*; Zhou et al. *ibid.* and Ensser and Fleckenstein, *ibid.* Overall semaphorins share the greatest degree of homology within the semaphorin domain, between, 25-93%, with a greater degree of divergence in all other 5 regions and domains suggesting distinct roles for various sub-groups within the semaphorin family. The viral semaphorins are the most diverse, sharing only 25% identity with vertebrate semaphorins. Between vertebrate and invertebrate semaphorins, the percent identity varies 10 between 30-40%.

Isolating and characterizing neurite growth cues would be of great value, for example, in modulating neurite growth and development; treatment of peripheral neuropathies; for use as therapeutics for the regeneration 15 of neurons following strokes, brain damage caused by head injuries and paralysis caused by spinal injuries; diagnosing neurological diseases and in treating neurodegenerative diseases such as multiple sclerosis, Alzheimer's disease and Parkinson's disease. The present 20 invention addresses these needs and others by providing novel semaphorins and related compositions and methods.

SUMMARY OF THE INVENTION

The present invention provides a novel semaphorin 25 polypeptide and related compositions and methods.

Within one aspect is provided an isolated polypeptide comprising a sequence of amino acids that is 96% identical to the amino acid residues as shown in SEQ ID NO:2 from Phe, residue 20, through Arg, residue 321, 30 wherein said polypeptide is a semaphorin. Within one embodiment any difference between said amino acid sequence of said isolated polypeptide and said corresponding amino acid sequence of SEQ ID NO:2 is due to a conservative amino acid substitution. Within another embodiment the 35 polypeptide comprises amino acid residues 20 through 321 of SEQ ID NO:2. Within another embodiment the polypeptide further comprises a heterologous affinity tag or binding

domain. Within a related embodiment the affinity tag is a Glu-Glu tag.

The invention also provides an isolated polypeptide having the amino acid sequence of SEQ ID NO:2.

5 The invention further provides an isolated polypeptide comprising the amino acid sequence of amino acid residue 1 to amino acid residue 19 of SEQ ID NO:2.

10 Additionally, the invention provides an isolated polypeptide consisting of a semaphorin domain, a transmembrane domain and a cytoplasmic domain, wherein said semaphorin domain comprises the amino acid sequence of SEQ ID NO:2.

15 Within another aspect is provided a fusion protein consisting essentially of a first portion and a second portion joined by a peptide bond, said first portion comprising a polypeptide wherein said polypeptide comprises a sequence of amino acids that is 96% identical to the amino acid residues as shown in SEQ ID NO:2 from Phe, residue 20, through Arg, residue 321; wherein said 20 polypeptide is a semaphorin; and said second portion comprising another polypeptide. Within one embodiment the second polypeptide is a transmembrane domain of a semaphorin protein. Within a related embodiment the second polypeptide further comprises a cytoplasmic domain of a 25 semaphorin protein.

30 The invention also provides a fusion protein comprising a secretory signal sequence having the amino acid sequence of amino acid residues 1 to 19 of SEQ ID NO:2, wherein said secretory signal sequence is operably linked to an additional polypeptide.

35 Within another aspect of the invention is provided an isolated polynucleotide selected from the group consisting of: a) a polynucleotide encoding a polypeptide comprising a sequence of amino acids that is 96% identical to the amino acid residues as shown in SEQ ID NO:2 from Phe, residue 20, through Arg, residue 321; b) degenerate polynucleotide sequences of a); and c) a polynucleotide

sequence complementary to a) or b); wherein said polynucleotide encodes a polypeptide which is a semaphorin. Within one embodiment the polypeptide comprises residues 20 through 321 of SEQ ID NO:2. Within another embodiment the 5 protein comprises residues 1 through 321 of SEQ ID NO:2.

Within another aspect is provided an expression vector comprising the following operably linked elements: a transcription promoter; a DNA segment encoding a polypeptide as described herein; and a transcription 10 terminator. Within one embodiment the expression vector further comprises secretory signal sequence operably linked to said DNA segment. Within yet another embodiment the expression vector further comprises an affinity tag operably linked to said DNA segment.

15 The invention also provided a cultured cell containing an expression vector as described herein, wherein said cell expresses said DNA segment. Within one embodiment the expression vector further comprises a secretory signal sequence operably linked to said DNA 20 segment and the cell secretes said polypeptide. Within yet another embodiment the expression vector further comprises an affinity tag operably linked to said DNA segment.

The invention also provides a method of making a semaphorin protein comprising: (a) culturing a host cell 25 containing an expression vector as described herein; and (b) recovering said protein encoded by said DNA segment. Within another embodiment the expression vector further comprises a secretory signal sequence operably linked to said DNA segment, the cell secretes the protein into a 30 culture medium, and the protein is recovered from the medium. Within a related embodiment the expression vector further comprises an affinity tag operably linked to said DNA segment.

Within a further aspect, the invention provides 35 an antibody that specifically binds to a polypeptide as described herein. Within one embodiment the antibody is selected from the group consisting of: a) polyclonal

antibody; b) murine monoclonal antibody; c) humanized antibody derived from b); and d) human monoclonal antibody. Within one embodiment the antibody fragment is selected from the group consisting of F(ab'), F(ab), Fab', Fab, Fv, 5 scFv, and minimal recognition unit. Within another embodiment is provided an anti-idiotype antibody that specifically binds to the antibody described herein. Within another embodiment is provided a binding protein that specifically binds to a polypeptide as described 10 herein.

Within another aspect the invention provides a pharmaceutical composition comprising a polypeptide as described herein, in combination with a pharmaceutically acceptable vehicle.

15 These and other aspects of the invention will become evident upon reference to the following detailed description and the attached drawing.

BRIEF DESCRIPTION OF THE DRAWING

20

Figure 1 shows an alignment of human (SEQ ID NO:2) and mouse (SEQ ID NO:4) ZSMF-3 with mouse semaphorin F (SEQ ID NO:14) and mouse semaphorin G (SEQ ID NO:13). The starting methionine, signal sequence and beginning of 25 the mature peptide are noted as are the conserved cysteine residues.

Figure 2 shows an alignment of human ZSMF-3 (SEQ ID NO:1) with murine Sema IVa (SEQ ID NO:37).

30 DETAILED DESCRIPTION OF THE INVENTION

Prior to setting forth the invention, it may be helpful to an understanding thereof to set forth definitions of certain terms to be used hereinafter:

Allelic variant : Any of two or more alternative 35 forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in phenotypic polymorphism within populations. Gene mutations can be silent (i.e., no change in the

encoded polypeptide), or may encode polypeptides having altered amino acid sequence. The term "allelic variant" is also used herein to denote a protein encoded by an allelic variant of a gene. Also included are the same protein from 5 the same species which differs from a reference amino acid sequence due to allelic variation. Allelic variation refers to naturally occurring differences among individuals in genes encoding a given protein.

Amino-terminal and carboxyl-terminal: are used 10 herein to denote positions within polypeptides and proteins. Where the context allows, these terms are used with reference to a particular sequence or portion of a polypeptide or protein to denote proximity or relative position. For example, a certain sequence positioned 15 carboxyl-terminal to a reference sequence within a protein is located proximal to the carboxyl terminus of the reference sequence, but is not necessarily at the carboxyl terminus of the complete protein.

Complement/anti-complement pair: Denotes non-20 identical moieties that form a non-covalently associated, stable pair under appropriate conditions. For instance, biotin and avidin (or streptavidin) are prototypical members of a complement/anti-complement pair. Other exemplary complement/anti-complement pairs include 25 receptor/ligand pairs, antibody/antigen (or hapten or epitope) pairs, sense/antisense polynucleotide pairs, and the like. Where subsequent dissociation of the complement/anti-complement pair is desirable, the complement/anti-complement pair preferably has a binding 30 affinity of $<10^{-9}$ M.

Complements of polynucleotide molecules: Denotes polynucleotide molecules having a complementary base sequence and reverse orientation as compared to a reference sequence. For example, the sequence 5' ATGCACGGG 3' is 35 complementary to 5' CCCGTGCAT 3'.

Degenerate: As applied to a nucleotide sequence such as a probe or primer, denotes a sequence of

nucleotides that includes one or more degenerate codons (as compared to a reference polynucleotide molecule that encodes a polypeptide). Degenerate codons contain different triplets of nucleotides, but encode the same 5 amino acid residue (i.e., GAU and GAC triplets each encode Asp).

Expression vector: A DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of interest operably linked to additional segments that 10 provide for its transcription. Such additional segments may include promoter and terminator sequences, and optionally one or more origins of replication, one or more selectable markers, an enhancer, a polyadenylation signal, and the like. Expression vectors are generally derived 15 from plasmid or viral DNA, or may contain elements of both.

Isolated: When applied to a polynucleotide molecule, the term "isolated" indicates that the molecule is removed from its natural genetic milieu, and is thus free of other extraneous or unwanted coding sequences, and 20 is in a form suitable for use within genetically engineered protein production systems. Such isolated molecules are those that are separated from their natural environment and include cDNA and genomic clones. Isolated DNA molecules of the present invention are free of other genes with which 25 they are ordinarily associated, and may include naturally occurring 5' and 3' untranslated regions such as promoters and terminators. The identification of associated regions will be evident to one of ordinary skill in the art (see for example, Dynan and Tijan, Nature 316:774-78, 1985).

30 Operably linked: As applied to nucleotide segments, the term "operably linked" indicates that the segments are arranged so that they function in concert for their intended purposes, e.g., transcription initiates in the promoter and proceeds through the coding segment to the 35 terminator.

Ortholog or species homolog: A polypeptide or protein obtained from one species that has homology to an

analogous polypeptide or protein from a different species. Sequence differences among orthologs are the result of speciation.

5 Polynucleotide: denotes a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. Polynucleotides include RNA and DNA, and may be isolated from natural sources, synthesized *in vitro*, or prepared from a combination of natural and synthetic molecules.

10 Promoter: Denotes a portion of a gene containing DNA sequences that provide for the binding of RNA polymerase and initiation of transcription. Promoter sequences are commonly, but not always, found in the 5' non-coding regions of genes.

15 Receptor: A cell-associated protein, or a polypeptide subunit of such protein, that binds to a bioactive molecule (the "ligand") and mediates the effect of the ligand on the cell. Binding of ligand to receptor results in a change in the receptor (and, in some cases, 20 receptor multimerization, i.e., association of identical or different receptor subunits) that causes interactions between the effector domain(s) of the receptor and other molecule(s) in the cell. These interactions in turn lead to alterations in the metabolism of the cell. Metabolic 25 events that are linked to receptor-ligand interactions include gene transcription, phosphorylation, dephosphorylation, cell proliferation, increases in cyclic AMP production, mobilization of cellular calcium, mobilization of membrane lipids, cell adhesion, hydrolysis 30 of inositol lipids and hydrolysis of phospholipids.

Secretory signal sequence: A DNA sequence that encodes a polypeptide (a "secretory peptide") that, as a component of a larger polypeptide, directs the larger polypeptide through a secretory pathway of a cell in which 35 it is synthesized. The larger polypeptide is commonly cleaved to remove the secretory peptide during transit through the secretory pathway.

Soluble receptor/soluble ligand: A receptor/ligand polypeptide that is not bound to a cell membrane. Many cell-surface receptors/ligands have naturally occurring, soluble counterparts that are produced by proteolysis or translated from alternatively spliced mRNAs. They may also be produced by cleavage or by recombinant methods. Soluble receptors are most commonly ligand-binding receptor polypeptides and soluble ligands are most commonly receptor-binding polypeptides that lack transmembrane and cytoplasmic domains. Soluble receptors/ligands can comprise additional amino acid residues, such as affinity tags that provide for purification of the polypeptide or provide sites for attachment of the polypeptide to a substrate. Receptor/ligand polypeptides are said to be substantially free of transmembrane and intracellular polypeptide segments when they lack sufficient portions of these segments to provide membrane anchoring or signal transduction, respectively.

Splice variant: As used herein, denote alternative forms of RNA transcribed from a gene. Splice variation arises naturally through use of alternative splicing sites within a transcribed RNA molecule, or less commonly between separately transcribed RNA molecules, and may result in several mRNAs transcribed from the same gene. Splice variants may encode polypeptides having altered amino acid sequence. The term splice variant is also used herein to denote a protein encoded by a splice variant of an mRNA transcribed from a gene.

Molecular weights and lengths of polymers determined by imprecise analytical methods (e.g., gel electrophoresis) will be understood to be approximate values. When such a value is expressed as "about" X or "approximately" X, the stated value of X will be understood to be accurate to $\pm 10\%$.

The present invention is based in part upon the discovery of a novel human member of the semaphorin family.

designated "ZSMF-3". The human ZSMF-3 nucleotide sequence is represented in SEQ ID NO:1 and the deduced amino acid sequence in SEQ ID NO:2. A murine homologue is represented in SEQ ID NO:3 and the deduced amino acid sequence in SEQ 5 ID NO:4. Murine and human ZSMF-3 share 90% nucleotide identity and 95% amino acid identity.

The novel human ZSMF-3 semaphorin proteins and polypeptides encoded by polynucleotides of the present invention were initially identified by querying an EST 10 (Expressed Sequence Tag) database for sequences homologous to conserved motifs within the semaphorin family. Sequence analysis of the deduced amino acid sequence as represented in SEQ ID NO:2 indicate the presence of a 19 amino acid residue signal sequence (amino acid residues 1-19 of SEQ ID 15 NO:2) followed by at least 302 amino acid residues of a semaphorin domain (amino acid residues 20-321 of SEQ ID NO:2), a stop codon followed by a poly A region. Within the semaphorin domain are 7 conserved cysteine residues (amino acid residues 107, 117, 135, 144, 165, 258 and 283 20 of SEQ ID NO:2). Those skilled in the art will recognize that these domain boundaries are approximate, and are based on alignments with known proteins and predictions of protein folding.

The sequence of human and murine ZSMF-3, as 25 represented in SEQ ID NOS:2 and 4, are homologous to murine Sema IVa identified by Zhou et al., (Mol. Cell. Neurosci. 9: 26-41, 1997. Sema IVa has been described as having a 16 amino acid residue signal sequence, a 648 amino acid residue extracellular region containing a 500 amino acid 30 residue semaphorin domain. Within the semaphorin domain are 13 conserved cysteine residues and 7 N-linked glycosylation sites. The extracellular region is followed by a 24 amino acid transmembrane domain and a 216 amino acid residue cytoplasmic domain having one protein kinase C 35 site.

This information suggests that the human ZSMF-3 protein encoded by the nucleotide sequence described in SEQ

ID NO:1 represents an alternative splicing product or "splice variant" of a corresponding full length human ZSMF-3 protein. Splice variants of a full length DNA sequence include DNA sequences that result from mature RNA molecules 5 created by known eukaryotic RNA splicing processes wherein the intron sequence is removed and exon sequence is joined. A well defined splice site which begins at nucleotide 963 of SEQ ID NO:1, a poly A region (nucleotide 1033-1038 of SEQ ID NO:1) followed by a poly A tail (beginning at 10 nucleotide 1054 of SEQ ID NO:1) confirm that this sequence indeed represents a viable splice variant. Northern blot analysis indicates a predominant transcript in stomach tissue of 2 kb which agrees with the splice variant size. Two faint transcripts of 4 kb and 3 kb were detected in 15 spinal cord and adrenal gland which would indicate the existence of a full length ZSMF-3 protein.

Zhou et al. (ibid.) suggest that their Sema IVa, represents a new class of mammalian semaphorins which have structural homology to the insect Sema I proteins. 20 Although murine Sema IVa shares 25-36% amino acid identity with murine Sema III within the semaphorin domain, Sema IVa has a transmembrane domain and a cytoplasmic domain, although much larger, as are seen in the insect Sema I proteins. Previously described transmembrane containing 25 vertebrate semaphorins also have regions of Ig homology or transponder repeats which are not present in murine Sema IVa. It is possible that human ZSMF-3 is a human homolog of murine Sema IVa and a member of this new semaphorin family. A final determination can only be made once it is 30 determined what sequence lies C terminal to the ZSMF-3 semaphorin domain. Over a similar area in the semaphorin domain, murine ZSMF-3 (SEQ ID NO:4) and human ZSMF-3 (SEQ ID NO:2) share 100% and 95% identity respectively, with murine SEMA IVa described by Zhou et al., ibid. Over a 35 similar area, human ZSMF-3 shares 31% nucleotide identity with *Vaccinia* virus semaphorin Sema IV, 36% identity with *Drosophila* semaphorin II (D Sema II), 39% identity with

Tribolium confusum semaphorin I (T Sema I), Human semaphorin CD100, and chicken collapsin, 42% identity with *Drosophila* semaphorin I (D Sema I) and grasshopper semaphorin I (G Sema I), 44% identity with murine 5 semaphorin Sem G, 45% identity with murine semaphorin SemF, and 61% identity with murine semaphorin VIb. Figure 1 shows an alignment of human and mouse ZSMF-3 with murine semaphorins Sem F and Sem G. Figure 2 shows an alignment of human ZSMF-3 with murine SEMA IVa. The truncated 10 semaphorin domain of ZSMF-3 may suggest that it also shares an association with the viral semaphorin Sema IV, which has a truncated semaphorin domain of about 400 amino acids. An alternatively spliced 3.5 kb human semaphorin, SEMA3F, was described by Xiang et al. (Genomics 32:39-48, 1996). The 15 splice resulted in the addition of 96 bases to the message and retained the correct reading frame. Such features set human ZSMF-3 apart from previously described semaphorins and may be characteristic qualities which further define a new sub-group within the semaphorin family.

20 The DNA sequence shown in SEQ ID NO:1 or portions thereof can be used as probes or primers to obtain the corresponding full length ZSMF-3 DNA from cells or libraries (including cDNA and genomic libraries) from human tissues. Methods for isolating polynucleotide sequences 25 are known in the art. One example for obtaining the ZSMF-3 sequence uses mRNA obtained from a tissue or cell types that are likely to express the protein. Suitable sources of mRNA can be identified by probing Northern blots with probes designed from the sequences disclosed herein. 30 Suitable tissue sources include human spinal cord, adrenal gland, stomach and brain. A library is then prepared from mRNA of a positive tissue or cell line. The cDNA can then be isolated by a variety of methods, such as by probing with the sequence disclosed in SEQ ID NO:1 or some portion 35 thereof. A cDNA can also be cloned using the polymerase chain reaction, or PCR (Mullis, U.S. Patent 4,683,202), using primers designed from the sequences disclosed herein

and templates derived from appropriate tissues as discussed above. Additional oligonucleotides may be generated from the PCR derived sequence and SEQ ID NO:1, as needed, to obtain a complete full length cDNA encoding a ZSMF-3 polypeptide of about 890 amino acid residues. More preferably such full length ZSMF-3 polypeptides would have about 700-750 amino acid residues. Such a full length ZSMF-3 molecule would include a semaphorin domain of approximately 500 amino acid residues, a transmembrane domain of approximately 24 amino acid residues and a cytoplasmic domain of approximately 216 amino acid residues. Other methods such as 3' RACE using oligo sequences from SEQ ID NO:1 and templates derived from spinal cord, adrenal gland, brain or stomach tissue could 15 also be used.

Northern blot analysis of various human tissues was performed using a 200 bp human DNA probe (SEQ ID NO: 6). A 2 kb transcript was detected corresponding to ZSMF-3 in stomach tissue. The transcript size agrees with the 20 predicted size of the ZSMF-3 protein as disclosed in SEQ ID NO:2. Two faint transcripts of 4 kb and 3 kb were detected in spinal cord and adrenal gland tissue.

Chromosomal localization results showed that ZSMF-3 maps 384.88 cR_3000 from the top of the human 25 chromosome 5 linkage group on the WICGR radiation hybrid map. The proximal and distal framework markers were D5S492 and WI-3394, respectively. The use of surrounding markers positions ZSMF-3 in the 5q23.1-q23.2 region on the integrated LDB chromosome 5 map. Among the genes mapping 30 to this region are a neurogenic differentiation factor, NEUROD3, (Tamimi et al., *Genomics* 40:355-7, 1997); type IV calmodulin-dependent protein kinase, CAMK4, (Sikela et al., *Genomics* 4:21-7, 1989); genes associated with Charcot- 35 Marie-Tooth disease, a demyelinating peripheral motor and sensory neuropathy LeGuern et al., *Hum. Molec. Genet.* 5:1685-8, 1996) and ileal lipid-binding protein (Oelkers and Dawson, *Biochim. Biophys. Acta* 1257:199-202, 1995).

A degenerate polynucleotide sequence that encompasses all polynucleotides that encode the ZSMF-3 polypeptides of SEQ ID NOS: 2 or 4, amino acid residues 1-321, are disclosed in SEQ ID NO:5 (human sequence) and SEQ 5 ID NO:17 (mouse sequence). Thus, ZSMF-3 polypeptide-encoding polynucleotides ranging from nucleotide 1-963 of SEQ ID NO:5 and nucleotides 1-963 of SEQ ID NO:17 are contemplated by the present invention. Also contemplated by the present invention are fragments as described above 10 with respect to SEQ ID NO:1, which are formed from analogous regions of SEQ ID NO:5, wherein nucleotides 1-963 of SEQ ID NO:1 correspond to nucleotides 57-963 of SEQ ID NO:5. The nucleotide base codes in SEQ ID NOS:5 and 17 are summarized in Table 1.

TABLE 1

Base Code	Nucleotide	Base Code	Nucleotide Complement
A	A	T	T
C	C	G	G
G	G	C	C
T	T	A	A
R	A G	Y	C T
Y	C T	R	A G
M	A C	K	G T
K	G T	M	A C
S	C G	S	C G
W	A T	W	A T
H	A C T	D	A G T
B	C G T	V	A C G
V	A C G	B	C G T
D	A G T	H	A C T
N	A C G T	N	A C G T

The degenerate codons used in SEQ ID NOS:5 and 17, encompassing all possible codons for a given amino acid, are set forth in Table 2.

5

TABLE 2

Three Letter Code	One Letter Code	Synonymous Codons	Degenerate Codon
Cys	C	TGC TGT	TGY
Ser	S	AGC AGT TCA TCC TCG TCT	WSN
Thr	T	ACA ACC ACG ACT	ACN
Pro	P	CCA CCC CCG CCT	CCN
Ala	A	GCA GCC GCG GCT	GCN
Gly	G	GGA GGC GGG GGT	GGN
Asn	N	AAC AAT	AAY
Asp	D	GAC GAT	GAY
Glu	E	GAA GAG	GAR
Gln	Q	CAA CAG	CAR
His	H	CAC CAT	CAY
Arg	R	AGA AGG CGA CGC CGG CGT	MGN
Lys	K	AAA AAG	AAR
Met	M	ATG	ATG
Ile	I	ATA ATC ATT	ATH
Leu	L	CTA CTC CTG CTT TTA TTG	YTN
Val	V	GTA GTC GTG GTT	GTN
Phe	F	TTC TTT	TTY
Tyr	Y	TAC TAT	TAY
Trp	W	TGG	TGG
Ter	.	TAA TAG TGA	TRR
Asn Asp	B		RAY
Glu Gln	Z		SAR
Any	X		NNN

One of ordinary skill in the art will appreciate that some ambiguity is introduced in determining a degenerate codon, representative of all possible codons encoding each amino acid. For example, the degenerate 5 codon for serine (WSN) can, in some circumstances, encode arginine (AGR), and the degenerate codon for arginine (MGN) can, in some circumstances, encode serine (AGY). A similar relationship exists between codons encoding phenylalanine and leucine. Thus, some polynucleotides encompassed by the 10 degenerate sequence may encode variant amino acid sequences, but one of ordinary skill in the art can easily identify such variant sequences by reference to the amino acid sequence of SEQ ID NO:2. Such variant sequences can be tested for functionality as disclosed herein.

One of ordinary skill in the art will also appreciate that different species can exhibit "preferential codon usage." In general, see, Grantham, et al., Nuc. Acids Res. 8:1893-912, 1980; Haas, et al. Curr. Biol. 6:315-24, 1996; Wain-Hobson, et al., Gene 13:355-64, 1981; 20 Grosjean and Fiers, Gene 18:199-209, 1982; Holm, Nuc. Acids Res. 14:3075-87, 1986; Ikemura, J. Mol. Biol. 158:573-97, 1982. As used herein, the term "preferential codon usage" or "preferential codons" is a term of art referring to protein translation codons that are most frequently used in 25 cells of a certain species, thus favoring one or a few representatives of the possible codons encoding each amino acid (See Table 2). For example, the amino acid threonine (Thr) may be encoded by ACA, ACC, ACG, or ACT, but in mammalian cells ACC is the most commonly used codon; in other species, for example, insect cells, yeast, viruses or 30 bacteria, different Thr codons may be preferential. Preferential codons for a particular species can be introduced into the polynucleotides of the present invention by a variety of methods known in the art. 35 Introduction of preferential codon sequences into recombinant DNA can, for example, enhance production of the protein by making protein translation more efficient within

a particular cell type or species. Therefore, the degenerate codon sequences disclosed in SEQ ID NOS:5 and 17 serve as templates for optimizing expression of polynucleotides in various cell types and species commonly used in the art and disclosed herein. Sequences containing preferential codons can be tested and optimized for expression in various species, and tested for functionality as disclosed herein.

Within preferred embodiments of the invention, isolated polynucleotides will hybridize to similar sized regions of SEQ ID NO:1, or to sequences complementary thereto, under stringent conditions. In general, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typical stringent conditions are those in which the salt concentration is up to about 0.03 M at pH 7 and the temperature is at least about 60°C. As previously noted, the isolated polynucleotides of the present invention include DNA and RNA. Methods for isolating DNA and RNA are well known in the art. It is generally preferred to isolate RNA from stomach tissue, although DNA can also be prepared using RNA from other tissues or isolated as genomic DNA. Total RNA can be prepared using guanidine HCl extraction followed by isolation by centrifugation in a CsCl gradient (Chirgwin et al., Biochemistry 18:52-94, 1979). Poly (A)⁺ RNA is prepared from total RNA using the method of Aviv and Leder (Proc. Natl. Acad. Sci. USA 69:1408-12, 1972). Complementary DNA (cDNA) is prepared from poly(A)⁺ RNA using known methods. Polynucleotides encoding ZSMF-3 polypeptides are then identified and isolated by, for example, hybridization or PCR.

Those skilled in the art will recognize that the sequence disclosed in SEQ ID NO:1 represents a single allele of human ZSMF-3 and the sequence disclosed in SEQ ID

NO:3 represents a single allele of murine ZSMF-3, and that allelic variation and alternative splicing are expected to occur. Allelic variants and splice variants of this sequence can be cloned by probing cDNA or genomic libraries from different individuals or tissues according to standard procedures known in the art. Allelic variants of the DNA sequence shown in SEQ ID NOS:1 or 3, including those containing silent mutations and those in which mutations result in amino acid sequence changes, are within the scope of the present invention, as are proteins which are allelic variants of SEQ ID NOS:2 or 4. cDNAs generated from alternatively spliced mRNAs, which retain the properties of the ZSMF-3 polypeptides are included within the scope of the present invention, as are polypeptides encoded by such cDNAs and mRNAs.

The present invention further provides counterpart ligands and polynucleotides from other species ("species orthologs"). These species would include, but are not limited to, mammalian, avian, amphibian, reptile, fish, insect and other vertebrate and invertebrate species. Of particular interest are ZSMF-3 polypeptides from other mammalian species, including porcine, ovine, bovine, canine, feline, equine, and other primate polypeptides. Species orthologs of human ZSMF-3 can be cloned using information and compositions provided by the present invention in combination with conventional cloning techniques. For example, a cDNA can be cloned using mRNA obtained from a tissue or cell type that expresses the ligand. Suitable sources of mRNA can be identified by probing Northern blots with probes designed from the sequences disclosed herein. A library is then prepared from mRNA of a positive tissue or cell line. A ligand-encoding cDNA can then be isolated by a variety of methods, such as by probing with a complete or partial human cDNA or with one or more sets of degenerate probes based on the disclosed sequence. A cDNA can also be cloned by PCR, using primers designed from the sequences disclosed herein.

Within an additional method, the cDNA library can be used to transform or transfect host cells, and expression of the cDNA of interest can be detected with an antibody to the ligand. Similar techniques can also be applied to the 5 isolation of genomic clones.

The present invention also provides isolated polypeptides and proteins that are substantially homologous to the polypeptides and proteins of SEQ ID NO:2 and their species orthologs. By "isolated" is meant a protein or 10 polypeptide that is found in a condition other than its native environment, such as apart from blood and animal tissue. In a preferred form, the isolated protein or polypeptide is substantially free of other proteins or polypeptides, particularly other proteins or polypeptides 15 of animal origin. It is preferred to provide the proteins or polypeptides in a highly purified form, i.e. greater than 95% pure, more preferably greater than 99% pure. The term "substantially homologous" is used herein to denote proteins or polypeptides having 50%, preferably 60%, more 20 preferably at least 80%, sequence identity to the sequence shown in SEQ ID NO:2 or its species orthologs. Such proteins or polypeptides will more preferably be at least 90% identical, and most preferably 95% or more identical to SEQ ID NO:2 or its species orthologs. Particularly 25 preferred are polypeptides which are 96%, 97%, 98% or 99% identical to the sequence shown in SEQ ID NO:2 or its species orthologs. Within one embodiment, any difference between the amino acid sequence of the polypeptide and the corresponding amino acid sequence of SEQ ID NO:2 is due to 30 a conservative amino acid substitution. Percent sequence identity is determined by conventional methods. See, for example, Altschul et al., Bull. Math. Bio. 48: 603-16, 1986 and Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA 89:10915-9, 1992. Briefly, two amino acid sequences are 35 aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "blosum 62" scoring matrix of Henikoff and Henikoff

(ibid.) as shown in Table 3 (amino acids are indicated by the standard one-letter codes). The percent identity is then calculated as:

5

$$\frac{\text{Total number of identical matches}}{[\text{length of the longer sequence plus the number of gaps introduced into the longer sequence in order to align the two sequences}]} \times 100$$

Table 3

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
5	A	4																		
	R	-1	5																	
	N	-2	0	6																
	D	-2	-2	1	6															
	C	0	-3	-3	-3	9														
	Q	-1	1	0	0	-3	5													
10	E	-1	0	0	2	-4	2	5												
	G	0	-2	0	-1	-3	-2	-2	6											
	H	-2	0	1	-1	-3	0	0	-2	8										
	I	-1	-3	-3	-1	-3	-3	-4	-3	4										
	L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4								
	K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5							
	M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5						
	F	-2	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
	P	-1	-2	-1	-3	-1	-1	-2	-2	-3	-1	-2	-4	7						
	S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4			
	T	0	-1	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5			
	W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11	
	Y	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
20	V	0	-3	-3	-1	-2	-2	-3	-3	1	-2	1	-1	-2	-2	0	-3	-1	4	

Sequence identity of polynucleotide molecules is determined by similar methods using a ratio as disclosed above.

Substantially homologous proteins and polypeptides are characterized as having one or more amino acid substitutions, deletions or additions. These changes are preferably of a minor nature, that is conservative amino acid substitutions (see Table 4) and other substitutions that do not significantly affect the folding or activity of the protein or polypeptide; small deletions, typically of one to about 30 amino acids; and small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue, a small linker peptide of up to about 20-25 residues, or a small extension that facilitates purification (an affinity tag), such as a poly-histidine tract, Glu-Glu affinity tags, protein A (Nilsson et al., EMBO J. 4:1075, 1985; Nilsson et al., Methods Enzymol. 198:3, 1991), glutathione S transferase (Smith and Johnson, Gene 67:31, 1988), maltose binding protein (Kellerman and Ferenci, Meth. Enzymol. 90:459-63, 1982; Guan et al., Gene 67:21-30, 1987), thioredoxin, ubiquitin, cellulose binding protein, T7 polymerase, an immunoglobulin heavy chain constant region or other antigenic epitope or binding domain. See, in general Ford et al., Protein Expression and Purification 2: 95-107, 1991. DNAs encoding affinity tags are available from commercial suppliers (e.g., Pharmacia Biotech, Piscataway, NJ).

Table 4

Conservative amino acid substitutions

5

	Basic:	arginine
		lysine
		histidine
10	Acidic:	glutamic acid
		aspartic acid
	Polar:	glutamine
		asparagine
15	Hydrophobic:	leucine
		isoleucine
		valine
	Aromatic:	phenylalanine
		tryptophan
		tyrosine
20	Small:	glycine
		alanine
		serine
		threonine
		methionine

25

In addition to the 20 standard amino acids, non-standard amino acids (such as 4-hydroxyproline, 6-N-methyl lysine, 2-aminoisobutyric acid, isovaline and α -methyl serine) may be substituted for amino acid residues of ZSMF-3. A limited number of non-conservative amino acids, amino acids that are not encoded by the genetic code, and unnatural amino acids may be substituted for ZSMF-3 amino acid residues. "Unnatural amino acids" have been modified after protein synthesis, and/or have a chemical structure in their side chain(s) different from that of the standard amino acids. Unnatural amino acids can be chemically synthesized, or preferably, are commercially available, and include pipecolic acid, thiazolidine carboxylic acid,

dehydroproline, 3- and 4-methylproline, and 3,3-dimethylproline.

The proteins of the present invention can also comprise non-naturally occurring amino acid residues. Non-naturally occurring amino acids include, without limitation, trans-3-methylproline, 2,4-methanoproline, cis-4-hydroxyproline, trans-4-hydroxyproline, N-methyl-glycine, allo-threonine, methylthreonine, hydroxyethyl-cysteine, hydroxyethylhomocysteine, nitroglutamine, homo-glutamine, piperolic acid, tert-leucine, norvaline, 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, and 4-fluorophenylalanine. Several methods are known in the art for incorporating non-naturally occurring amino acid residues into proteins. For example, an *in vitro* system can be employed wherein nonsense mutations are suppressed using chemically aminoacylated suppressor tRNAs. Methods for synthesizing amino acids and aminoacylating tRNA are known in the art. Transcription and translation of plasmids containing nonsense mutations is carried out in a cell free system comprising an *E. coli* S30 extract and commercially available enzymes and other reagents. Proteins are purified by chromatography. See, for example, Robertson et al., *J. Am. Chem. Soc.* 113:2722, 1991; Ellman et al., *Methods Enzymol.* 202:301, 1991; Chung et al., *Science* 259:806-9, 1993; and Chung et al., *Proc. Natl. Acad. Sci. USA* 90:10145-9, 1993). In a second method, translation is carried out in *Xenopus* oocytes by microinjection of mutated mRNA and chemically aminoacylated suppressor tRNAs (Turcatti et al., *J. Biol. Chem.* 271:19991-8, 1996). Within a third method, *E. coli* cells are cultured in the absence of a natural amino acid that is to be replaced (e.g., phenylalanine) and in the presence of the desired non-naturally occurring amino acid(s) (e.g., 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, or 4-fluorophenylalanine). The non-naturally occurring amino acid is incorporated into the protein in place of its natural counterpart. See, Koide et al., *Biochem.* 33:7470-

6, 1994. Naturally occurring amino acid residues can be converted to non-naturally occurring species by *in vitro* chemical modification. Chemical modification can be combined with site-directed mutagenesis to further expand 5 the range of substitutions (Wynn and Richards, Protein Sci. 2:395-3, 1993).

Essential amino acids in the polypeptides of the present invention can be identified according to procedures known in the art, such as site-directed mutagenesis or 10 alanine-scanning mutagenesis (Cunningham and Wells, Science 244: 1081-5, 1989; Bass et al., Proc. Natl. Acad. Sci. USA 88:4498-502, 1991). In the latter technique, single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for 15 biological activity (e.g., collapase assay) to identify amino acid residues that are critical to the activity of the molecule. Sites of ligand-receptor interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic 20 resonance, crystallography, electron diffraction, or photoaffinity; in conjunction with mutation of putative contact site amino acids. See, for example, de Vos et al., Science 255:306-12, 1992; Smith et al., J. Mol. Biol. 224:899-904, 1992; Wlodaver et al., FEBS Lett. 309:59-64, 25 1992. The identities of essential amino acids can also be inferred from analysis of homologies with related polypeptides.

Multiple amino acid substitutions can be made and tested using known methods of mutagenesis and screening, 30 such as those disclosed by Reidhaar-Olson and Sauer (Science 241:53-7, 1988) or Bowie and Sauer (Proc. Natl. Acad. Sci. USA 86:2152-6, 1989). Briefly, these authors disclose methods for simultaneously randomizing two or more positions in a polypeptide, selecting for functional 35 polypeptide, and then sequencing the mutagenized polypeptides to determine the spectrum of allowable substitutions at each position. Other methods that can be

used include phage display (e.g., Lowman et al., Biochem. 30:10832-7, 1991; Ladner et al., U.S. Patent No. 5,223,409; Huse, WIPO Publication WO 92/06204) and region-directed mutagenesis (Derbyshire et al., Gene 46:145, 1986; Ner et al., DNA 7:127, 1988).

Mutagenesis methods as disclosed above can be combined with high-throughput screening methods to detect activity of cloned, mutagenized polypeptides in host cells. Mutagenized DNA molecules that encode active ligands or portions thereof (e.g., receptor-binding fragments) can be recovered from the host cells and rapidly sequenced using modern equipment. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide of interest, and can be applied to polypeptides of unknown structure.

Using the methods discussed above, one of ordinary skill in the art can identify and/or prepare a variety of polypeptides that are substantially homologous to amino acid residues 20-321 of SEQ ID NO:2 or allelic variants thereof and retain the properties of the wild-type protein. Such polypeptides may include additional amino acids from the semaphorin, transmembrane and/or cytoplasmic domains; affinity tags; and the like. Such polypeptides may also include additional polypeptide segments as generally disclosed above.

The semaphorin polypeptides of the present invention, including full-length polypeptides, fragments (e.g., receptor-binding fragments, growth cone directing fragments, immune response provoking fragments), and fusion polypeptides, can be produced in genetically engineered host cells according to conventional techniques. Suitable host cells are those cell types that can be transformed or transfected with exogenous DNA and grown in culture, and include bacteria, fungal cells, and cultured higher eukaryotic cells. Eukaryotic cells, particularly cultured cells of multicellular organisms, are preferred. Techniques for manipulating cloned DNA molecules and

introducing exogenous DNA into a variety of host cells are disclosed by Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, NY, 1989; and Ausubel et al., eds., Current Protocols in Molecular Biology, John Wiley and Sons, Inc., NY, 1987.

In general, a DNA sequence encoding a ZSMF-3 polypeptide is operably linked to other genetic elements required for its expression, generally including a transcription promoter and terminator, within an expression vector. The vector will also commonly contain one or more selectable markers and one or more origins of replication, although those skilled in the art will recognize that within certain systems selectable markers may be provided on separate vectors, and replication of the exogenous DNA may be provided by integration into the host cell genome. Selection of promoters, terminators, selectable markers, vectors and other elements is a matter of routine design within the level of ordinary skill in the art. Many such elements are described in the literature and are available through commercial suppliers.

To direct a ZSMF-3 polypeptide into the secretory pathway of a host cell, a secretory signal sequence (also known as a leader sequence, prepro sequence or pre sequence) is provided in the expression vector. The secretory signal sequence may be that of the ZSMF-3 polypeptide, or may be derived from another secreted protein (e.g., t-PA) or synthesized *de novo*. The secretory signal sequence is joined to the ZSMF-3 DNA sequence in the correct reading frame. Secretory signal sequences are commonly positioned 5' to the DNA sequence encoding the polypeptide of interest, although certain signal sequences may be positioned elsewhere in the DNA sequence of interest (see, e.g., Welch et al., U.S. Patent No. 5,037,743; Holland et al., U.S. Patent No. 5,143,830).

Alternatively, the secretory signal sequence contained in the polypeptides of the present invention is used to direct other polypeptides into the secretory

pathway. The present invention provides for such fusion polypeptides. A signal fusion polypeptide can be made wherein a secretory signal sequence derived from amino acid residues 1-19 of SEQ ID NOS:2 or 4 is operably linked to 5 another polypeptide using methods known in the art and disclosed herein. The secretory signal sequence contained in the fusion polypeptides of the present invention is preferably fused amino-terminally to an additional peptide to direct the additional peptide into the secretory 10 pathway. Such constructs have numerous applications known in the art. For example, these novel secretory signal sequence fusion constructs can direct the secretion of an active component of a normally non-secreted protein. Such fusions may be used *in vivo* or *in vitro* to direct peptides 15 through the secretory pathway.

Cultured mammalian cells are preferred hosts within the present invention. Methods for introducing exogenous DNA into mammalian host cells include calcium phosphate-mediated transfection (Wigler et al., Cell 20 14:725, 1978; Corsaro and Pearson, Somatic Cell Genetics 7:603, 1981; Graham and Van der Eb, Virology 52:456, 1973), electroporation (Neumann et al., EMBO J. 1:841-5, 1982), DEAE-dextran mediated transfection (Ausubel et al., ibid.), and liposome-mediated transfection (Hawley-Nelson et al., 25 Focus 15:73, 1993; Ciccarone et al., Focus 15:80, 1993). The production of recombinant polypeptides in cultured mammalian cells is disclosed, for example, by Levinson et al., U.S. Patent No. 4,713,339; Hagen et al., U.S. Patent No. 4,784,950; Palmiter et al., U.S. Patent No. 4,579,821; 30 and Ringold, U.S. Patent No. 4,656,134. Suitable cultured mammalian cells include the COS-1 (ATCC No. CRL 1650), COS-7 (ATCC No. CRL 1651), BHK (ATCC No. CRL 1632), BHK 570 (ATCC No. CRL 10314), 293 (ATCC No. CRL 1573; Graham et al., J. Gen. Virol. 36:59-72, 1977) and Chinese hamster 35 ovary (e.g., CHO-K1; ATCC No. CCL 61) cell lines. Additional suitable cell lines are known in the art and available from public depositories such as the American

Type Culture Collection, Rockville, Maryland. In general, strong transcription promoters are preferred, such as promoters from SV-40 or cytomegalovirus. See, e.g., U.S. Patent No. 4,956,288. Other suitable promoters include 5 those from metallothionein genes (U.S. Patent Nos. 4,579,821 and 4,601,978) and the adenovirus major late promoter.

Drug selection is generally used to select for cultured mammalian cells into which foreign DNA has been 10 inserted. Such cells are commonly referred to as "transfectants". Cells that have been cultured in the presence of the selective agent and are able to pass the gene of interest to their progeny are referred to as "stable transfectants." A preferred selectable marker is a 15 gene encoding resistance to the antibiotic neomycin. Selection is carried out in the presence of a neomycin-type drug, such as G-418 or the like. Selection systems may also be used to increase the expression level of the gene of interest, a process referred to as "amplification." 20 Amplification is carried out by culturing transfectants in the presence of a low level of the selective agent and then increasing the amount of selective agent to select for cells that produce high levels of the products of the introduced genes. A preferred amplifiable selectable 25 marker is dihydrofolate reductase, which confers resistance to methotrexate. Other drug resistance genes (e.g., hygromycin resistance, multi-drug resistance, puromycin acetyltransferase) can also be used. Alternative markers that introduce an altered phenotype, such as green 30 fluorescent protein, or cell surface proteins such as CD4, CD8, Class I MHC, placental alkaline phosphatase may be used to sort transfected cells from untransfected cells by such means as FACS sorting or magnetic bead separation technology.

35 Other higher eukaryotic cells can also be used as hosts, including insect cells, plant cells and avian cells. Transformation of insect cells and production of foreign

polypeptides therein is disclosed by Guarino et al., U.S. Patent No. 5,162,222; Bang et al., U.S. Patent No. 4,775,624; and WIPO publication WO 94/06463. The use of *Agrobacterium rhizogenes* as a vector for expressing genes in plant cells has been reviewed by Sinkar et al., J. Biosci. (Bangalore) 11:47-58, 1987. Insect cells can be infected with recombinant baculovirus vectors, which are commonly derived from *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV). DNA encoding the polypeptide of interest is inserted into the viral genome in place of the polyhedrin gene coding sequence by homologous recombination in cells infected with intact, wild-type AcMNPV and transfected with a transfer vector comprising the cloned gene operably linked to polyhedrin gene promoter, terminator, and flanking sequences. The resulting recombinant virus is used to infect host cells, typically a cell line derived from the fall armyworm, *Spodoptera frugiperda*. See, in general, Glick and Pasternak, Molecular Biotechnology: Principles and Applications of Recombinant DNA, ASM Press, Washington, D.C., 1994.

Fungal cells, including yeast cells, and particularly cells of the genus *Saccharomyces*, can also be used within the present invention, such as for producing fragments or polypeptide fusions. Methods for transforming yeast cells with exogenous DNA and producing recombinant polypeptides therefrom are disclosed by, for example, Kawasaki, U.S. Patent No. 4,599,311; Kawasaki et al., U.S. Patent No. 4,931,373; Brake, U.S. Patent No. 4,870,008; Welch et al., U.S. Patent No. 5,037,743; and Murray et al., U.S. Patent No. 4,845,075. Transformed cells are selected by phenotype determined by the selectable marker, commonly drug resistance or the ability to grow in the absence of a particular nutrient (e.g., leucine). A preferred vector system for use in yeast is the POT1 vector system disclosed by Kawasaki et al. (U.S. Patent No. 4,931,373), which allows transformed cells to be selected by growth in

glucose-containing media. Suitable promoters and terminators for use in yeast include those from glycolytic enzyme genes (see, e.g., Kawasaki, U.S. Patent No. 4,599,311; Kingsman et al., U.S. Patent No. 4,615,974; and 5 Bitter, U.S. Patent No. 4,977,092) and alcohol dehydrogenase genes. See also U.S. Patents Nos. 4,990,446; 5,063,154; 5,139,936 and 4,661,454. Transformation systems for other yeasts, including *Hansenula polymorpha*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, 10 *Kluyveromyces fragilis*, *Ustilago maydis*, *Pichia pastoris*, *Pichia methanolica*, *Pichia guillermondii* and *Candida maltosa* are known in the art. See, for example, Gleeson et al., J. Gen. Microbiol. 132:3459-65, 1986; and Cregg, U.S. Patent No. 4,882,279. *Aspergillus* cells may be utilized 15 according to the methods of McKnight et al., U.S. Patent No. 4,935,349. Methods for transforming *Acremonium chrysogenum* are disclosed by Sumino et al., U.S. Patent No. 5,162,228. Methods for transforming *Neurospora* are disclosed by Lambowitz, U.S. Patent No. 4,486,533.

20 For example, the use of *Pichia methanolica* as host for the production of recombinant proteins is disclosed by Raymond, U.S. Patent No. 5,716,808, Raymond, U.S. Patent No. 5,736,383, Raymond et al., Yeast 14:11-23, 1998, and in WIPO Publication Nos. WO 97/17450, WO 25 97/17451, WO 98/02536, and WO 98/02565.

Prokaryotic host cells, including strains of the bacteria *Escherichia coli*, *Bacillus* and other genera are also useful host cells within the present invention. Techniques for transforming these hosts and expressing 30 foreign DNA sequences cloned therein are well known in the art (see, e.g., Sambrook et al., ibid.). When expressing a ZSMF-3 polypeptide in bacteria such as *E. coli*, the polypeptide may be retained in the cytoplasm, typically as insoluble granules, or may be directed to the periplasmic 35 space by a bacterial secretion sequence. In the former case, the cells are lysed, and the granules are recovered and denatured using, for example, guanidine isothiocyanate

or urea. The denatured polypeptide can then be refolded and dimerized by diluting the denaturant, such as by dialysis against a solution of urea and a combination of reduced and oxidized glutathione, followed by dialysis 5 against a buffered saline solution. In the latter case, the polypeptide can be recovered from the periplasmic space in a soluble and functional form by disrupting the cells (by, for example, sonication or osmotic shock) to release the contents of the periplasmic space and recovering the 10 protein, thereby obviating the need for denaturation and refolding.

Transformed or transfected host cells are cultured according to conventional procedures in a culture medium containing nutrients and other components required 15 for the growth of the chosen host cells. A variety of suitable media, including defined media and complex media, are known in the art and generally include a carbon source, a nitrogen source, essential amino acids, vitamins and minerals. Media may also contain such components as growth 20 factors or serum, as required. The growth medium will generally select for cells containing the exogenously added DNA by, for example, drug selection or deficiency in an essential nutrient which is complemented by the selectable marker carried on the expression vector or co-transfected 25 into the host cell.

The invention provides soluble ZSMF-3 ligand polypeptides. Soluble ZSMF-3 polypeptides can be prepared by expressing a DNA encoding amino acid residues 20-321 of a human ZSMF-3 polypeptide (SEQ ID NO:2). In one 30 embodiment, it is preferred that these soluble, extracellular receptor-binding domain polypeptides be prepared in a form substantially free of transmembrane and intracellular polypeptide segments. For example, the C-terminus of the ligand polypeptide may be at residue 321 of 35 SEQ ID NO:2 or the corresponding region of an allelic variant or at the C-terminal end of a semaphorin domain. To direct the export from the host cell, the ligand DNA is

linked to a second DNA segment encoding a secretory peptide, such as a t-PA secretory peptide. To facilitate purification of the secreted receptor domain, an N- or C-terminal extension, such as a poly-histidine tag, a Glu-Glu 5 affinity tag, substance P, FLAG™ peptide (Hopp et al., *Biotechnology* 6:1204-10, 1988; available from Eastman Kodak Co., New Haven, CT) or another polypeptide or protein for which an antibody or other specific binding agent is available, can be fused to the ligand polypeptide.

Expressed recombinant ZSMF-3 polypeptides (or chimeric or fused ZSMF-3 polypeptides) can be purified using fractionation and/or conventional purification methods and media. Ammonium sulfate precipitation and acid or chaotropic extraction may be used for fractionation of 10 samples. Exemplary purification steps may include hydroxyapatite, size exclusion, FPLC and reverse-phase high performance liquid chromatography. Suitable anion exchange media include derivatized dextrans, agarose, cellulose, polyacrylamide, specialty silicas, and the like. PEI, 15 DEAE, QAE and Q derivatives are preferred, with DEAE Fast-Flow Sepharose (Pharmacia, Piscataway, NJ) being particularly preferred. Exemplary chromatographic media include those media derivatized with phenyl, butyl, or octyl groups, such as Phenyl-Sepharose FF (Pharmacia), 20 Toyopearl butyl 650 (Toso Haas, Montgomeryville, PA), Octyl-Sepharose (Pharmacia) and the like; or polyacrylic resins, such as Amberchrom CG 71 (Toso Haas) and the like. Suitable solid supports include glass beads, silica-based 25 resins, cellulosic resins, agarose beads, cross-linked agarose beads, polystyrene beads, cross-linked 30 polyacrylamide resins and the like that are insoluble under the conditions in which they are to be used. These supports may be modified with reactive groups that allow attachment of proteins by amino groups, carboxyl groups, 35 sulfhydryl groups, hydroxyl groups and/or carbohydrate moieties. Examples of coupling chemistries include cyanogen bromide activation, N-hydroxysuccinimide

activation, epoxide activation, sulfhydryl activation, hydrazide activation, and carboxyl and amino derivatives for carbodiimide coupling chemistries. These and other solid media are well known and widely used in the art, and 5 are available from commercial suppliers. Methods for binding receptor polypeptides to support media are well known in the art. Selection of a particular method is a matter of routine design and is determined in part by the properties of the chosen support. See, for example, 10 Affinity Chromatography: Principles & Methods, Pharmacia LKB Biotechnology, Uppsala, Sweden, 1988.

The polypeptides of the present invention can be isolated by exploitation of their physical properties. For example, immobilized metal ion adsorption (IMAC) 15 chromatography can be used to purify histidine-rich proteins. Briefly, a gel is first charged with divalent metal ions to form a chelate (Sulkowski, Trends in Biochem. 3:1-7, 1985). Histidine-rich proteins will be adsorbed to this matrix with differing affinities, depending upon the 20 metal ion used, and will be eluted by competitive elution, lowering the pH, or use of strong chelating agents. Other methods of purification include purification of glycosylated proteins by lectin affinity chromatography and ion exchange chromatography (Methods in Enzymol., Vol. 182, 25 "Guide to Protein Purification", M. Deutscher, (ed.), Acad. Press, San Diego, 1990, pp.529-39). Alternatively, a fusion of the polypeptide of interest and an affinity tag (e.g., polyhistidine, maltose-binding protein, an immunoglobulin domain) may be constructed to facilitate 30 purification.

It is preferred to purify the protein to >80% purity, more preferably to >90% purity, even more preferably >95%, and particularly preferred is a pharmaceutically pure state, that is greater than 99.9% 35 pure with respect to contaminating macromolecules, particularly other proteins and nucleic acids, and free of infectious and pyrogenic agents. Preferably, a purified

protein is substantially free of other proteins, particularly other proteins of animal origin.

ZSMF-3 polypeptides or fragments thereof may also be prepared through chemical synthesis. ZSMF-3 5 polypeptides may be monomers or multimers; glycosylated or non-glycosylated; pegylated or non-pegylated; and may or may not include an initial methionine amino acid residue.

Polypeptides containing the receptor-binding region of the ligand can be used for purification of 10 receptor. The ligand polypeptide is immobilized on a solid support, such as beads of agarose, cross-linked agarose, glass, cellulosic resins, silica-based resins, polystyrene, cross-linked polyacrylamide, or like materials that are 15 stable under the conditions of use. Methods for linking polypeptides to solid supports are known in the art, and include amine chemistry, cyanogen bromide activation, N-hydroxysuccinimide activation, epoxide activation, sulfhydryl activation, and hydrazide activation. The 20 resulting media will generally be configured in the form of a column, and fluids containing receptors are passed through the column one or more times to allow receptor to bind to the ligand polypeptide. The receptor is then eluted using changes in salt concentration, chaotropic agents ($MnCl_2$), or pH to disrupt ligand-receptor binding.

25 Soluble ZSMF-3 or ZSMF-3 fusion proteins are used, for example, to identify the ZSMF-3 receptor. Using labeled soluble ZSMF-3, cells expressing the receptor are identified by fluorescence immunocytometry or immunohistochemistry. The soluble receptor is useful in 30 studying the distribution of receptor on tissues or specific cell lineages, and to provide insight into receptor/ligand biology.

ZSMF-3 fusion proteins consist essentially of a first portion and a second portion joined by a peptide 35 bond. The first portion comprising a polypeptide wherein the polypeptide comprises a sequence of amino acid residues selected from the group consisting of: a) the sequence of

amino acids as shown in SEQ ID NO:2 from Phe, residue 20, through Arg, residue 321; b) a sequence of amino acids that is 96% identical to the amino acid residues as shown in SEQ ID NO:2 from Phe, residue 20, through Arg, residue 321; and 5 c) allelic variants of a) or b); wherein said polypeptide is a semaphorin. The second portion comprises another polypeptide. The second portion polypeptide may be another polypeptide such as a semaphorin or a polypeptide fragment of a semaphorin protein, such as a transmembrane domain or 10 cytoplasmic domain.

ZSMF-3 proteins, agonists and antagonists may be used to modulate neurite growth and development and demarcate nervous system structures. As such, ZSMF-3 and antagonists would be useful as a treatment of peripheral 15 neuropathies by increasing spinal cord and sensory neurite outgrowth. Such an antagonist could be part of a therapeutic treatment for the regeneration of neurite outgrowths following strokes, brain damage caused by head injuries and paralysis caused by spinal injuries. 20 Application may also be made in treating neurodegenerative diseases such as multiple sclerosis, Alzheimer's disease and Parkinson's disease. Application may also be made in mediating development and innervation pattern of stomach tissue.

25 Semaphorins have also been associated with mediation of other non-neuronal biological functions. ZSMF-3 polypeptides may find use in mediating immune response, in particular ZSMF-3 polypeptides and ZSMF-3 agonists would be useful as immunosuppressants. It has 30 been suggested that viral proteins play a role in modulating immune response against viral infection by either binding to a cytokine or preventing activation of a cytokine. It is speculated that viral semaphorin, Sema IV, plays the role of a natural immunosuppressor to reduce the 35 immune response (Kolodkin et al., *ibid.*). ZSMF-3 shares structural homology with sema IV and may have similar biological activity.

Other non-viral semaphorins have also been associated with the immune system. Human semaphorin E, which is homologous to viral cytokine inhibiting proteins contains conserved regions of amino acid residues that have 5 been found in the viral semaphorins. Semaphorin E was found to be upregulated in rheumatoid synovial fibroblastoid cells which suggests that it may have a role as a regulator of inflammatory processes and an involvement in the development of rheumatoid arthritis (Mangasser-10 Stephan et al., Biochem. Biophys. Res. Comm. 234:153-6, 1997). Semaphorin G (Sema G) is expressed on lymphocytes and may play a role in the immune system (Furuyama et al., J. Biol. Chem. 271:33376-81, 1996). The figure shows the homology ZSMF-3 shares with sema G. Semaphorin CD100 15 promotes B-cell growth and aggregation and may be involved in lymphocyte activation (Hall et al., Proc. Natl. Acad. Sci. USA 93: 11780-5, 1996).

Semaphorins, such as Semaphorin III, have a role 20 in the development of bones and heart by acting as a restraining signal during organ development (Behar et al., Nature 383:525-8, 1996). G-Sema I and collapsin are hypothesized to act *in vivo* as repulsive or inhibitory molecules that prevent neighboring ventral motorneurons from innervating extra thoracic muscle. In other 25 situations, G-Sema I and collapsin may also act as an attractive agent to promote innervation (Kolodkin et al., Cell 75:1389-99, 1993). The truncated human ZSMF-3 may play a similar role in the stomach.

The activity of semaphorin polypeptides, 30 agonists, antagonists and antibodies of the present invention can be measured, and compounds screened to identify agonists and antagonists, using a variety of assays, in particular assays that measure axon guidance and growth. Of particular interest are assays that indicate 35 changes in neuron growth patterns, see for example, Hastings, WIPO Patent Application No:97/29189 and Walter et al., Development 101:685-96, 1987. Assays to measure the

effects of semaphorins on neuron growth are well known in the art. For example, the C assay (see for example, Raper and Kapfhammer, Neuron 4:21-9, 1990 and Luo et al., Cell 75:217-27, 1993), can be used to determine collapsing 5 activity semaphorins on growing neurons. Other methods which assess semaphorin induced inhibition of neurite extension or divert such extension are also known, see Goodman, Annu. Rev. Neurosci. 19:341-77, 1996. Conditioned media from cells expressing a semaphorin, semaphorin 10 agonist or semaphorin antagonist, or aggregates of such cells, can be placed in a gel matrix near suitable neural cells, such as dorsal root ganglia (DRG) or sympathetic ganglia explants, which have been cocultured with nerve growth factor. Compared to control cells, semaphorin- 15 induced changes in neuron growth can be measured (see for example, Messersmith et al., Neuron 14:949-59, 1995; Puschel et al., Neuron 14:941-8, 1995). Likewise neurite outgrowth can be measured using neuronal cell suspensions grown in the presence of molecules of the present invention 20 see for example, O'Shea et al., Neuron 7:231-7, 1991 and DeFreitas et al., Neuron 15:333-43, 1995.

Also available are assay systems that use a ligand-binding receptor (or an antibody, one member of a complement/anti-complement pair) or a binding fragment 25 thereof, and a commercially available biosensor instrument (BIAcore™, Pharmacia Biosensor, Piscataway, NJ). Such receptor, antibody, member of a complement/anti-complement pair or fragment is immobilized onto the surface of a receptor chip. Use of this instrument is disclosed by 30 Karlsson, J. Immunol. Methods 145:229-40, 1991 and Cunningham and Wells, J. Mol. Biol. 234:554-63, 1993. A receptor, antibody, member or fragment is covalently attached, using amine or sulfhydryl chemistry, to dextran fibers that are attached to gold film within the flow cell. 35 A test sample is passed through the cell. If a ligand, epitope, or opposite member of the complement/anti-complement pair is present in the sample, it will bind to

the immobilized receptor, antibody or member, respectively, causing a change in the refractive index of the medium, which is detected as a change in surface plasmon resonance of the gold film. This system allows the determination of 5 on- and off-rates, from which binding affinity can be calculated, and assessment of stoichiometry of binding. Ligand-binding receptor polypeptides can also be used within other assay systems known in the art. Such systems include Scatchard analysis for determination of binding 10 affinity (see, Scatchard, Ann. NY Acad. Sci. 51: 660-72, 1949) and calorimetric assays (Cunningham et al., Science 253:545-8, 1991; Cunningham et al., Science 245:821-5, 1991).

15 *In vitro* and *in vivo* response to soluble ZSMF-3 can also be measured using cultured cells or by administering molecules of the claimed invention to the appropriate animal model. Proteins of the present invention may also be assayed using viral delivery systems. Exemplary viruses for this purpose include adenovirus, 20 herpesvirus, vaccinia virus and adeno-associated virus (AAV). Adenovirus, a double-stranded DNA virus, is currently the best studied gene transfer vector for delivery of heterologous nucleic acid (for a review, see Becker et al., Meth. Cell Biol. 43:161-89, 1994; and 25 Douglas and Curiel, Science & Medicine 4:44-53, 1997). A preferred method is that of He et al., Proc. Natl. Acad. Sci. USA 95:2509-14, 1998 and is described in more detail below.

30 The adenovirus system can also be used for protein production *in vitro*. By culturing adenovirus-infected non-293 cells under conditions where the cells are not rapidly dividing, the cells can produce proteins for extended periods of time. For instance, BHK cells are grown to confluence in cell factories, then exposed to the 35 adenoviral vector encoding the secreted protein of interest. The cells are then grown under serum-free conditions, which allows infected cells to survive for

several weeks without significant cell division. Alternatively, adenovirus vector infected 293S cells can be grown in suspension culture at relatively high cell density to produce significant amounts of protein (see A. Garnier 5 et al., Cytotechnol. 15:145-55, 1994). With either protocol, an expressed, secreted heterologous protein can be repeatedly isolated from the cell culture supernatant. Within the infected 293S cell production protocol, non-secreted proteins may also be effectively obtained.

10 ZSMF-3 polypeptides can also be used to prepare antibodies that specifically bind to ZSMF-3 epitopes, peptides or polypeptides. Methods for preparing polyclonal and monoclonal antibodies are well known in the art (see, for example, Sambrook et al., Molecular Cloning: A 15 Laboratory Manual, Second Edition, Cold Spring Harbor, NY, 1989; and Hurrell, J. G. R., Ed., Monoclonal Hybridoma Antibodies: Techniques and Applications, CRC Press, Inc., Boca Raton, FL, 1982). As would be evident to one of ordinary skill in the art, polyclonal antibodies can be 20 generated from a variety of warm-blooded animals, such as horses, cows, goats, sheep, dogs, chickens, rabbits, mice, and rats.

25 The immunogenicity of a ZSMF-3 polypeptide may be increased through the use of an adjuvant, such as alum (aluminum hydroxide) or Freund's complete or incomplete adjuvant. Polypeptides useful for immunization also include fusion polypeptides, such as fusions of ZSMF-3 or a portion thereof with an immunoglobulin polypeptide or with maltose binding protein. The polypeptide immunogen may be 30 a full-length molecule or a portion thereof. If the polypeptide portion is "hapten-like", such portion may be advantageously joined or linked to a macromolecular carrier (such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA) or tetanus toxoid) for immunization.

35 As used herein, the term "antibodies" includes polyclonal antibodies, affinity-purified polyclonal antibodies, monoclonal antibodies, and antigen-binding

fragments thereof, such as $F(ab')_2$ and Fab proteolytic fragments. Genetically engineered intact antibodies or fragments, such as chimeric antibodies, Fv fragments, single chain antibodies (scFv), single complementarity-determining region peptides ("minimal recognition units") and the like, as well as synthetic antigen-binding peptides and polypeptides, are also included. Non-human antibodies may be humanized by grafting only non-human CDRs onto human framework and constant regions, or by incorporating the entire non-human variable domains (optionally "cloaking" them with a human-like surface by replacement of exposed residues, wherein the result is a "veeneered" antibody). In some instances, humanized antibodies may retain non-human residues within the human variable region framework domains to enhance proper binding characteristics. Through humanizing antibodies, biological half-life may be increased, and the potential for adverse immune reactions upon administration to humans is reduced. Humanized monoclonal antibodies directed against ZSMF-3 polypeptides could be used as a protein therapeutic, in particular for use as an immunotherapy. Alternative techniques for generating or selecting antibodies useful herein include *in vitro* exposure of lymphocytes to ZSMF-3 protein or peptide, and selection of antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled ZSMF-3 protein or peptide).

Antibodies are defined to be specifically binding if they bind to a ZSMF-3 polypeptide with a binding affinity (K_a) of 10^6 M^{-1} or greater, preferably 10^7 M^{-1} or greater, more preferably 10^8 M^{-1} or greater, and most preferably 10^9 M^{-1} or greater. The binding affinity of an antibody can be readily determined by one of ordinary skill in the art (for example, by Scatchard analysis).

Genes encoding polypeptides having potential ZSMF-3 binding domains can be obtained by screening random peptide libraries displayed on phage (phage display) or on bacteria, such as *E. coli*. Nucleotide sequences encoding

the polypeptides can be obtained in a number of ways, such as through random mutagenesis and random polynucleotide synthesis. These random peptide display libraries can be used to screen for peptides which interact with a known target which can be a protein or polypeptide, such as a ligand or receptor, a biological or synthetic macromolecule, or organic or inorganic substances. Techniques for creating and screening such random peptide display libraries are known in the art (Ladner et al., US Patent NO:5,223,409; Ladner et al., US Patent NO:4,946,778; 10 Ladner et al., US Patent NO:5,403,484 and Ladner et al., US Patent NO:5,571,698) and random peptide display libraries and kits for screening such libraries are available commercially, for instance from Clontech (Palo Alto, CA), 15 Invitrogen Inc. (San Diego, CA), New England Biolabs, Inc. (Beverly, MA) and Pharmacia LKB Biotechnology Inc. (Piscataway, NJ). Random peptide display libraries can be screened using the ZSMF-3 sequences disclosed herein to identify proteins which bind to ZSMF-3. These "binding 20 proteins" which interact with ZSMF-3 polypeptides may be used for tagging cells; for isolating homolog polypeptides by affinity purification; they can be directly or indirectly conjugated to drugs, toxins, radionuclides and the like.

25 A variety of assays known to those skilled in the art can be utilized to detect antibodies and binding proteins which specifically bind to ZSMF-3 proteins or peptides. Exemplary assays are described in detail in Antibodies: A Laboratory Manual, Harlow and Lane (Eds.), 30 Cold Spring Harbor Laboratory Press, 1988. Representative examples of such assays include: concurrent immunoelectrophoresis, radioimmunoassay, radioimmuno-precipitation, enzyme-linked immunosorbent assay (ELISA), dot blot or Western blot assay, inhibition or competition assay, and sandwich assay. In addition, antibodies can be screened for binding to wild-type versus mutant ZSMF-3 protein or peptide.

Antibodies and binding proteins to ZSMF-3 may be used for immunohistochemical tagging of cells that express human ZSMF-3, for example, to use in diagnostic assays; for isolating ZSMF-3 by affinity purification; for screening 5 expression libraries; for generating anti-idiotypic antibodies; and as neutralizing antibodies or as antagonists to block ZSMF-3 *in vitro* and *in vivo*. Suitable direct tags or labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent markers, 10 chemiluminescent markers, magnetic particles and the like; indirect tags or labels may feature use of biotin-avidin or other complement/anti-complement pairs as intermediates. Antibodies and binding proteins herein may also be directly or indirectly conjugated to drugs, toxins, radionuclides, 15 and the like, and these conjugates used for *in vivo* diagnostic or therapeutic applications.

The invention also provides isolated and purified ZSMF-3 polynucleotide probes. Such polynucleotide probes can be RNA or DNA. DNA can be either cDNA or genomic DNA. 20 Polynucleotide probes are single or double-stranded DNA or RNA, generally synthetic oligonucleotides, but may be generated from cloned cDNA or genomic sequences and will generally comprise at least 16 nucleotides, more often from 17 nucleotides to 25 or more nucleotides, sometimes 40 to 25 60 nucleotides, and in some instances a substantial portion, domain or even the entire ZSMF-3 gene or cDNA. The synthetic oligonucleotides of the present invention have at least 80% identity to a representative ZSMF-3 DNA sequence (SEQ ID NOS:1 or 3) or its complements. Preferred regions from which to construct probes include the 5' and/or 3' coding sequences, receptor binding regions, semaphorin, extracellular, transmembrane and/or cytoplasmic domains and signal sequences, and the like. Techniques for developing polynucleotide probes and hybridization 30 35 techniques are known in the art, see for example, Ausubel et al., eds., Current Protocols in Molecular Biology, John Wiley and Sons, Inc., NY, 1991. For use as probes, the

molecules can be labeled to provide a detectable signal, such as with an enzyme, biotin, a radionuclide, fluorophore, chemiluminescer, paramagnetic particle and the like, which are commercially available from many sources, 5 such as Molecular Probes, Inc., Eugene, OR, and Amersham Corp., Arlington Heights, IL, using techniques that are well known in the art.

Such probes can also be used in hybridizations to detect the presence or quantify the amount of ZSMF-3 gene 10 or mRNA transcript in a sample. ZSMF-3 polynucleotide probes could be used to hybridize to DNA or RNA targets for diagnostic purposes, using such techniques such as fluorescent *in situ* hybridization (FISH) or immunohistochemistry. Polynucleotide probes can be used to 15 identify genes encoding ZSMF-3-like proteins. For example, ZSMF-3 polynucleotides can be used as primers and/or templates in PCR reactions to identify other novel members of the semaphorin family. Such probes can also be used to screen libraries for related sequences encoding novel 20 semaphorin receptors. Such screening would be carried out under conditions of low stringency which would allow identification of sequences which are substantially homologous, but not requiring complete homology to the 25 probe sequence. Such methods and conditions are well known in the art, see, for example, Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, NY, 1989. Such low stringency conditions could 30 include hybridization temperatures less than 42°C, formamide concentrations of less than 50% and moderate to low concentrations of salt. Libraries may be made of genomic DNA or cDNA. Polynucleotide probes are also useful 35 for Southern, Northern, or slot blots, colony and plaque hybridization and *in situ* hybridization. Mixtures of different ZSMF-3 polynucleotide probes can be prepared which would increase sensitivity or the detection of low copy number targets, in screening systems. Nucleic acid probes can be detectably labeled with radioisotopes such as

³²P or ³⁵S. Alternatively, ZSMF-3 RNA can be detected with a nonradioactive hybridization method (see, for example, Isaac (ed.), Protocols for Nucleic Acid Analysis by Nonradioactive Probes, Humana Press, Inc., 1993).

5 Typically, nonradioactive detection is achieved by enzymatic conversion of chromogenic or chemiluminescent substrates. Illustrative non-radioactive moieties include biotin, fluorescein, and digoxigenin.

Numerous diagnostic procedures take advantage of

10 the polymerase chain reaction (PCR) to increase sensitivity of detection methods. Standard techniques for performing PCR are well-known (see, generally, Mathew (ed.), Protocols in Human Molecular Genetics, Humana Press, Inc., 1991; White (ed.), PCR Protocols: Current Methods and

15 Applications, Humana Press, Inc., 1993; Cotter (ed.), Molecular Diagnosis of Cancer, Humana Press, Inc., 1996; Hanausek and Walaszek (eds.), Tumor Marker Protocols, Humana Press, Inc., 1998; Lo (ed.), Clinical Applications of PCR, Humana Press, Inc., 1998 and Meltzer (ed.), PCR in

20 Bioanalysis, Humana Press, Inc., 1998). PCR primers can be designed to amplify a sequence encoding a particular ZSMF-3 region, such as the semaphorin domain.

One variation of PCR for diagnostic assays is reverse transcriptase-PCR (RT-PCR). In the RT-PCR

25 technique, RNA is isolated from a biological sample, reverse transcribed to cDNA, and the cDNA is incubated with ZSMF-3 primers (see, for example, Wu et al. (eds.), "Rapid Isolation of Specific cDNAs or Genes by PCR," in Methods in Gene Biotechnology, pages 15-28, CRC Press, Inc. 1997).

30 PCR is then performed and the products are analyzed using standard techniques.

As an illustration, RNA is isolated from a biological sample using, for example, the guanidinium-thiocyanate cell lysis procedure described above.

35 Alternatively, a solid-phase technique can be used to isolate mRNA from a cell lysate. A reverse transcription reaction can be primed with the isolated RNA using random

oligonucleotides, short homopolymers of dT, or ZSMF-3 anti-sense oligomers. Oligo-dT primers offer the advantage that various mRNA nucleotide sequences are amplified that can provide control target sequences. ZSMF-3 sequences are 5 amplified by the polymerase chain reaction using two flanking oligonucleotide primers that are typically 20 bases in length.

PCR amplification products can be detected using a variety of approaches. For example, PCR products can be 10 fractionated by gel electrophoresis, and visualized by ethidium bromide staining. Alternatively, fractionated PCR products can be transferred to a membrane, hybridized with a detectably-labeled ZSMF-3 probe, and examined by autoradiography. Additional alternative approaches include 15 the use of digoxigenin-labeled deoxyribonucleic acid triphosphates to provide chemiluminescence detection, and the C-TRAK colorimetric assay.

Another approach is real time quantitative PCR (Perkin-Elmer Cetus, Norwalk, CT). A fluorogenic probe, 20 consisting of an oligonucleotide with both a reporter and a quencher dye attached, anneals specifically between the forward and reverse primers. Using the 5' endonuclease activity of Taq DNA polymerase, the reporter dye is separated from the quencher dye and a sequence-specific 25 signal is generated and increases as amplification increases. The fluorescence intensity can be continuously monitored and quantified during the PCR reaction.

Another approach for detection of ZSMF-3 expression is cycling probe technology (CPT), in which a 30 single-stranded DNA target binds with an excess of DNA-RNA-DNA chimeric probe to form a complex, the RNA portion is cleaved with RNase H, and the presence of cleaved chimeric probe is detected (see, for example, Beggs et al., J. Clin. Microbiol. 34:2985, 1996 and Bekkaoui et al., Biotechniques 35 20:240, 1996). Alternative methods for detection of ZSMF-3 sequences can utilize approaches such as nucleic acid sequence-based amplification (NASBA), cooperative

amplification of templates by cross-hybridization (CATCH), and the ligase chain reaction (LCR) (see, for example, Marshall et al., U.S. Patent No. 5,686,272 (1997), Dyer et al., J. Virol. Methods 60:161, 1996; Ehricht et al., Eur. J. Biochem. 243:358, 1997; and Chadwick et al., J. Virol. Methods 70:59, 1998). Other standard methods are known to those of skill in the art.

10 ZSMF-3 probes and primers can also be used to detect and to localize ZSMF-3 gene expression in tissue samples. Methods for such *in situ* hybridization are well-known to those of skill in the art (see, for example, Choo (ed.), In Situ Hybridization Protocols, Humana Press, Inc., 1994; Wu et al. (eds.), "Analysis of Cellular DNA or Abundance of mRNA by Radioactive *In Situ* Hybridization 15 (RISH)," in Methods in Gene Biotechnology, pages 259-278, CRC Press, Inc., 1997; and Wu et al. (eds.), "Localization of DNA or Abundance of mRNA by Fluorescence *In Situ* Hybridization (RISH)," in Methods in Gene Biotechnology, pages 279-289, CRC Press, Inc., 1997).

20 Various additional diagnostic approaches are well-known to those of skill in the art (see, for example, Mathew (ed.), Protocols in Human Molecular Genetics, Humana Press, Inc., 1991; Coleman and Tsongalis, Molecular Diagnostics, Humana Press, Inc., 1996; and Elles, Molecular 25 Diagnosis of Genetic Diseases, Humana Press, Inc., 1996).

30 In addition, such polynucleotide probes could be used to hybridize to counterpart sequences on individual chromosomes. Chromosomal identification and/or mapping of the ZSMF-3 gene could provide useful information about gene function and disease association. Many mapping techniques are available to one skilled in the art, for example, mapping somatic cell hybrids, and fluorescence *in situ* hybridization (FISH). A preferred method is radiation hybrid mapping. Radiation hybrid mapping is a somatic cell 35 genetic technique developed for constructing high-resolution, contiguous maps of mammalian chromosomes (Cox et al., Science 250:245-50, 1990). Partial or full

knowledge of a gene's sequence allows the designing of PCR primers suitable for use with chromosomal radiation hybrid mapping panels. Commercially available radiation hybrid mapping panels which cover the entire human genome, such as 5 the Stanford G3 RH Panel and the GeneBridge 4 RH Panel (Research Genetics, Inc., Huntsville, AL), are available. These panels enable rapid, PCR based, chromosomal localizations and ordering of genes, sequence-tagged sites (STSs), and other non-polymorphic- and polymorphic markers 10 within a region of interest. This includes establishing directly proportional physical distances between newly discovered genes of interest and previously mapped markers. The precise knowledge of a gene's position can be useful in a number of ways including: 1) determining if a sequence is 15 part of an existing contig and obtaining additional surrounding genetic sequences in various forms such as YAC-, BAC- or cDNA clones, 2) providing a possible candidate gene for an inheritable disease which shows linkage to the same chromosomal region, and 3) for cross-referencing model organisms such as mouse which may be 20 beneficial in helping to determine what function a particular gene might have.

Chromosomal localization can also be done using STSSs. An STS is a DNA sequence that is unique in the human 25 genome and can be used as a reference point for a particular chromosome or region of a chromosome. An STS can be defined by a pair of oligonucleotide primers that can be used in a polymerase chain reaction to specifically detect this site in the presence of all other genomic 30 sequences. Since STSSs are based solely on DNA sequence they can be completely described within a database, for example, Database of Sequence Tagged Sites (dbSTS), GenBank, (National Center for Biological Information, National Institutes of Health, Bethesda, MD 35 <http://www.ncbi.nlm.nih.gov>), they can be searched with a gene sequence of interest for the mapping data contained within these short genomic landmark STS sequences.

ZSMF-3 was mapped to the 5q23.1-q23.2 region on the integrated LDB chromosome 5 map as is described in more detail below.

The present invention also contemplates use of such chromosomal localization for diagnostic applications. Briefly, the ZSMF-3 gene, a probe comprising ZSMF-3 DNA or RNA or a subsequence thereof, can be used to determine if the ZSMF-3 gene is present on human chromosome 5 or if a mutation has occurred. Detectable chromosomal aberrations at the ZSMF-3 gene locus include, but are not limited to, aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements. Such aberrations can be detected using polynucleotides of the present invention by employing molecular genetic techniques, such as restriction fragment length polymorphism (RFLP) analysis, short tandem repeat (STR) analysis employing PCR techniques, and other genetic linkage analysis techniques known in the art (Sambrook et al., ibid.; Ausubel et. al., ibid.; Marian, Chest 108:255-65, 1995).

In general, these diagnostic methods comprise the steps of (a) obtaining a genetic sample from a patient; (b) incubating the genetic sample with a polynucleotide probe or primer as disclosed above, under conditions wherein the polynucleotide will hybridize to complementary polynucleotide sequence, to produce a first reaction product; and (iii) comparing the first reaction product to a control reaction product. A difference between the first reaction product and the control reaction product is indicative of a genetic abnormality in the patient. Genetic samples for use within the present invention include genomic DNA, cDNA, and RNA. The polynucleotide probe or primer can be RNA or DNA, and will comprise a portion of SEQ ID NO:1, the complement of SEQ ID NO:1, or an RNA equivalent thereof. Suitable assay methods in this regard include molecular genetic techniques known to those in the art, such as restriction fragment length

polymorphism (RFLP) analysis, short tandem repeat (STR) analysis employing PCR techniques, ligation chain reaction (Barany, PCR Methods and Applications 1:5-16, 1991), ribonuclease protection assays, and other genetic linkage analysis techniques known in the art (Sambrook et al., ibid.; Ausubel et. al., ibid.; Marian, Chest 108:255-65, 1995). Ribonuclease protection assays (see, e.g., Ausubel et al., ibid., ch. 4) comprise the hybridization of an RNA probe to a patient RNA sample, after which the reaction product (RNA-RNA hybrid) is exposed to RNase. Hybridized regions of the RNA are protected from digestion. Within PCR assays, a patient's genetic sample is incubated with a pair of polynucleotide primers, and the region between the primers is amplified and recovered. Changes in size or amount of recovered product are indicative of mutations in the patient. Another PCR-based technique that can be employed is single strand conformational polymorphism (SSCP) analysis (Hayashi, PCR Methods and Applications 1:34-8, 1991).

ZSMF-3 polypeptides and polynucleotides may be used within diagnostic systems. Antibodies or other agents that specifically bind to ZSMF-3 may be used to detect the presence of circulating ligand or receptor polypeptides. Such detection methods are well known in the art and include, for example, enzyme-linked immunosorbent assay (ELISA) and radioimmunoassay. Immunohistochemically labeled ZSMF-3 antibodies can be used to detect ZSMF-3 receptor and/or ligands in tissue samples and identify ZSMF-3 receptors. ZSMF-3 levels can also be monitored by such methods as RT-PCR, where ZSMF-3 mRNA can be detected and quantified. The information derived from such detection methods would provide insight into the significance of ZSMF-3 polypeptides in various diseases and biological processes, and as a would serve as diagnostic tools for diseases for which altered levels of ZSMF-3 are significant.

The ZSMF-3 polynucleotides and/or polypeptides disclosed herein can be useful as therapeutics, wherein ZSMF-3 agonists and antagonists could modulate one or more biological processes in cells, tissues and/or biological fluids. ZSMF-3 antagonists provided by the invention, bind to ZSMF-3 polypeptides or, alternatively, to a receptor to which ZSMF-3 polypeptides bind, thereby inhibiting or eliminating the function of ZSMF-3. Such ZSMF-3 antagonists would include antibodies; oligonucleotides which bind either to the ZSMF-3 polypeptide or to its ligand; natural or synthetic analogs of ZSMF-3 ligands which retain the ability to bind the receptor but do not result in either ligand or receptor signaling. Such analogs could be peptides or peptide-like compounds. Natural or synthetic small molecules which bind to ZSMF-3 polypeptides and prevent signaling are also contemplated as antagonists. As such, ZSMF-3 antagonists would be useful as therapeutics for treating certain disorders where blocking signal from either a ZSMF-3 receptor or ligand would be beneficial.

The invention also provides nucleic acid-based therapeutic treatment. If a mammal lacks or has a mutated ZSMF-3 gene, the ZSMF-3 gene can be introduced into the cells of the mammal. Using such methods, cells altered to express the nerve growth factor neurotrophin-3 (NT-3) were grafted to a rat model for spinal injury and stimulated axon regrowth at the lesion site and the rats thus treated recovered some ability to walk (Grill et al., J. Neuroscience 17:5560-72, 1997). In one embodiment, a gene encoding a ZSMF-3 polypeptide is introduced *in vivo* in a viral vector. Such vectors include an attenuated or defective DNA virus, such as but not limited to herpes simplex virus (HSV), papillomavirus, Epstein Barr virus (EBV), adenovirus, adeno-associated virus (AAV), and the like. Defective viruses, which entirely or almost entirely lack viral genes, are preferred. A defective virus is not infective after introduction into a cell. Use of defective

viral vectors allows for administration to cells in a specific, localized area, without concern that the vector can infect other cells. Examples of particular vectors include, but are not limited to, a defective herpes virus 1 (HSV1) vector (Kaplitt et al., Molec. Cell. Neurosci. 2:320-30, 1991), an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet et al. (J. Clin. Invest. 90:626-30, 1992), and a defective adeno-associated virus vector (Samulski et al., J. Virol. 61:3096-101, 1987; Samulski et al., J. Virol. 63:3822-8, 1989).

In another embodiment, the gene can be introduced in a retroviral vector, e.g., as described in Anderson et al., U.S. Patent No. 5,399,346; Mann et al., Cell 33:153, 1983; Temin et al., U.S. Patent No. 4,650,764; Temin et al., U.S. Patent No. 4,980,289; Markowitz et al., J. Virol. 62:1120, 1988; Temin et al., U.S. Patent No. 5,124,263; Dougherty et al., WIPO Publication WO 95/07358; and Kuo et al., Blood 82:845-52, 1993.

Alternatively, the vector can be introduced by lipofection *in vivo* using liposomes. Synthetic cationic lipids can be used to prepare liposomes for *in vivo* transfection of a gene encoding a marker (Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7, 1987; and Mackey et al., Proc. Natl. Acad. Sci. USA 85:8027-31, 1988). The use of lipofection to introduce exogenous genes into specific organs *in vivo* has certain practical advantages. Molecular targeting of liposomes to specific cells represents one area of benefit. It is clear that directing transfection to particular cells represents one area of benefit. It is clear that directing transfection to particular cell types would be particularly advantageous in a tissue with cellular heterogeneity, such as the pancreas, liver, kidney, and brain. Lipids may be chemically coupled to other molecules for the purpose of targeting. Targeted peptides, e.g., hormones or neurotransmitters, and proteins

such as antibodies, or non-peptide molecules could be coupled to liposomes chemically.

It is possible to remove the cells from the body and introduce the vector as a naked DNA plasmid and then 5 re-implant the transformed cells into the body. Naked DNA vector for gene therapy can be introduced into the desired host cells by methods known in the art, e.g., transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a 10 gene gun or use of a DNA vector transporter (see, for example, Wu et al., J. Biol. Chem. 267:963-7, 1992; Wu et al., J. Biol. Chem. 263:14621-4, 1988).

Another aspect of the present invention involves 15 antisense polynucleotide compositions that are complementary to a segment of the polynucleotide set forth in SEQ ID NO:1. Such synthetic antisense oligonucleotides are designed to bind to mRNA encoding ZSMF-3 polypeptides and to inhibit translation of such mRNA. Such antisense oligonucleotides are used to inhibit expression of ZSMF-3 20 polypeptide-encoding genes in cell culture or in a subject.

Transgenic mice, engineered to express the ZSMF-3 gene, and mice that exhibit a complete absence of ZSMF-3 gene function, referred to as "knockout mice" (Snouwaert et al., Science 257:1083, 1992), may also be generated (Lowell 25 et al., Nature 366:740-2, 1993). These mice may be employed to study the ZSMF-3 gene and the protein encoded thereby in an *in vivo* system.

The invention is further illustrated by the following non-limiting examples.

EXAMPLESExample 1
Identification of ZSMF-3

5

Novel ZSMF-3 encoding polynucleotides and polypeptides of the present invention were initially identified by querying an EST database for sequences homologous to conserved motifs within the semaphorin family. To identify the corresponding cDNA, a clone from which an identified EST was derived that was considered likely to contain the entire murine ZSMF-3 sequence was used for sequencing. Using a QIAwell 8 plasmid kit (Qiagen, Inc., Chatsworth, CA) according to manufacturer's instructions, a 5 ml overnight culture in LB + 50 µg/ml 10 ampicillin was prepared. The template was sequenced on an Applied Biosystems™ model 377 DNA sequencer (Perkin-Elmer Cetus, Norwalk, Ct.) using the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer Corp.) 15 according to the manufacturer's instructions. Oligonucleotides ZC694 (SEQ ID NO:7) and ZC976 (SEQ ID NO:8) to the T7 and SP6 promoters on the vector were used as sequencing primers. Oligonucleotides ZC14201 (SEQ ID NO:9) and ZC14202 (SEQ ID NO:10) were used to complete the 20 sequence from the clone. Sequencing reactions were carried out in a Hybaid OmniGene Temperature Cycling System 25 (National Labnet Co., Woodbridge, NY). SequencherTM 3.0 sequence analysis software (Gene Codes Corporation, Ann Arbor, MI) was used for data analysis. A 1047 bp sequence 30 was identified which encoded a semaphorin protein having a truncated semaphorin domain, but no stop codon was identified.

To identify the human ZSMF-3 cDNA, a clone from which the identified human EST was derived and considered 35 likely to contain the entire human ZSMF-3 sequence was used for sequencing. Using a QIAwell 8 plasmid kit as above an overnight culture was prepared. The template was sequenced

as above using oligos ZC694 (SEQ ID NO:7), ZC976 (SEQ ID NO:8) as sequencing primers. This clone contained only about 600 bp of 3' sequence, so 5' RACE (rapid amplification of cDNA ends) was done to identify the 5 remaining 5' portion. Using a fetal brain Marathon™ cDNA library as a template and oligos ZC15,523 (SEQ ID NO:18) to a conserved region of the mouse untranslated region and ZC14,095 (SEQ ID NO:11) to an internal sequence of the 3' human fragment, 5' RACE was carried out at 94°C, for 1.5 10 minutes, followed by 35 cycles at 94°C for 15 seconds, 56°C for 30 seconds and 72°C for 1 minute, followed by a 10 minute extension at 72°C. A band of approximately 700 bp fragment was resolved by gel electrophoresis. The 5' RACE fragment was ligated into the TA vector (Invitrogen Inc., 15 San Diego, CA) according to manufacturer's instructions. The sequence of the 5' fragment was confirmed by sequence analysis as described above.

Homologous recombination in yeast was used to create expression plasmids containing the polynucleotide 20 encoding ZSMF-3 for expression in mammalian and yeast cells. To construct the ZSMF-3/pCZF199 expression vectors the following DNA fragments were transformed into *S. cerevisiae*: Sna BI digested pCZR199 as an acceptor vector, the 3' ZSMF-3 Xho I-Not I restriction fragment, the 5' Eco 25 RI restriction fragment, and two, double stranded linker segment. The expression vector, pCZR199, has yeast replication elements, (CEN, ARS), the selectable marker, URA3, *E. coli* replication elements (e.g., AMP^R and ori), a blunt-ended cloning site, Sna BI, and adds either a N-terminal or C-terminal Glu-Glu tag. The vectors were used 30 to create ZSMF-3 polypeptides having either end of the expressed protein Glu-Glu tagged. The double stranded linker segments were prepared using PCR. The linkers served to join the vector to the insert fragments at both 35 the 5' and 3' ends. Two sets of linkers were prepared. One set of linkers joined the insert to a vector placing the Glu-Glu tag on the 5' end of the insert sequence. The

5' linker of this set was made from oligos ZC14,397 (SEQ ID NO:27), ZC14,396 (SEQ ID NO:28), ZC16,463 (SEQ ID NO:29) and ZC16,442 (SEQ ID NO:30). The 3' linker was made from oligos ZC14,455 (SEQ ID NO:31), ZC14,394 (SEQ ID NO:32), 5 ZC16,461 (SEQ ID NO:33) and ZC16,444 (SEQ ID NO:26). The second set of linkers was used to join the ZSMF-3 insert into a vector placing a 3' Glu-Glu tag. The 5' linker was made from oligos ZC14,454 (SEQ ID NO:21), ZC14,392 (SEQ ID NO:22), ZC16,443 (SEQ ID NO:19) and ZC16,462 (SEQ ID NO:20). The 3' linker was made from oligos ZC14,395 (SEQ ID NO:24), ZC14,393 (SEQ ID NO:23), ZC16,441 (SEQ ID NO:25) and ZC16,444 (SEQ ID NO:26). The oligos were joined using standard PCR reaction conditions and heated to 94°C for 1.5 minutes followed by 10 cycles at 94°C for 30 seconds, 50°C 10 minutes for 1 minute and 72°C for 1 minute, then a 10 minute 15 extension at 72°C.

The DNA fragments were added to 100 µl competent yeast (Genetic strain SF83890, Roffman et al., EMBO J. 8:2057-65, 1989) and electroporated. The yeast cells were 20 immediately diluted in 600 µl 1.5 M sorbitol and plated on Ura⁺ D plates and incubated at 30°C for 48 hours. Ura⁺ colonies were selected from both the N-terminally-tagged and C-terminally-tagged ZSMF-3 proteins and the DNA from 25 the resulting yeast colonies was extracted and transformed into *E. coli*. Individual clones harboring the correct expression construct were identified by PCR screening. DNA sequencing as described above using oligos ZC447 (SEQ ID NO:34), ZC16296 (SEQ ID NO:35) and ZC16,292 (SEQ ID NO:36) as sequencing primers confirmed that the desired 30 sequences had been enjoined with one another.

Large scale plasmid DNA is isolated from one or more correct clones from both the N- and C-terminally tagged ZSMF-3 sequences, the expression cassette liberated from the vector and transformed into yeast or *E. coli* for 35 large scale protein production.

Example 2
Tissue Distribution

Human Multiple Tissue Northern Blots (MTN I, MTN II, and MTN III; Clontech) were probed to determine the tissue distribution of human ZSMF-3 expression. An approximately 200 bp PCR derived probe (SEQ ID NO:6) was amplified from a human stomach derived MarathonTM-ready cDNA library. Oligonucleotide primers ZC14,095 (SEQ ID NO:11) and ZC14,096 (SEQ ID NO:12) were designed based on the EST sequence. The MarathonTM-ready cDNA library was prepared according to manufacturer's instructions (MarathonTM cDNA Amplification Kit; Clontech, Palo Alto, CA) using human stomach poly A+ RNA (Clontech). The probe was amplified in a polymerase chain reaction as follows: 1 cycle at 94°C for 1 minute 30 seconds; 35 cycles of 94°C for 15 seconds and 64°C for 30 seconds, followed by 1 cycle at 72°C for 10 minutes. The resulting DNA fragment was electrophoresed on a 2% low melt agarose gel (SEA PLAQUE GTG low melt agarose, FMC Corp., Rockland, ME), the fragment was purified using the QIAquickTM method (Qiagen, Chatsworth, CA), and the sequence was confirmed by sequence analysis. The probe was radioactively labeled using the random priming MULTIPRIME DNA labeling system (Amersham, Arlington Heights, IL), according to the manufacturer's specifications. The probe was purified using a NUCTRAP push column (Stratagene, La Jolla, CA). ExpressHybTM (Clontech) solution was used for prehybridization and as a hybridizing solution for the Northern blots. Hybridization took place overnight at 65°C using 2 x 10⁶ cpm/ml of labeled probe. The blots were then washed at 25°C in 2X SSC, 0.05% SDS for 1 hour, then at 50°C in 0.1X SSC, 0.1% SDS for 1 hour. A transcript of approximately 2 kb was seen in stomach.

A Human RNA Master Dot Blot (Clontech) that contained RNAs from various tissues that were normalized to

8 housekeeping genes was also probed and hybridized as described above. Expression was seen only in stomach.

5 Mouse Multiple Choice[®] Tissue Northern Blots (OriGene Technologies, Inc., Rockville, MD) were probed to determine the tissue distribution of mouse ZSMF-3 expression. The 200 bp stomach derived probe (SEQ ID NO:6) described above was used as a probe. Hybridization took place overnight at 65°C using 1 x 10⁶ cpm/ml of labeled probe. The blots were then washed at 25°C in 2X SSC, 0.1% 10 SDS for 1 hour, then at 50°C in 0.1X SSC, 0.1% SDS for 1 hour. A transcript of approximately 1.5 kb was seen in all tissues, stomach, small intestine, muscle, lung, skin, brain, heart, kidney, spleen, thymus and liver.

15 A mouse RNA Master BlotTM (Clontech) was probed as described above. Expression was detected only in epididymus.

Example 3

Chromosomal Assignment and Placement of ZSMF-3

20 ZSMF-3 was mapped to chromosome 5 using the commercially available GeneBridge 4 Radiation Hybrid Panel (Research Genetics, Inc., Huntsville, AL). The GeneBridge 4 Radiation Hybrid Panel contains PCRable DNAs from each of 25 93 radiation hybrid clones, plus two control DNAs (the HFL donor and the A23 recipient). A publicly available WWW server (<http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>) allows mapping relative to the Whitehead Institute/MIT Center for Genome Research's radiation hybrid 30 map of the human genome (the "WICGR" radiation hybrid map) which was constructed with the GeneBridge 4 Radiation Hybrid Panel.

For the mapping of ZSMF-3 with the GeneBridge 4 RH Panel, 20 µl reactions were set up in a 96-well 35 microtiter plate (Stratagene) and used in a RoboCycler Gradient 96 thermal cycler (Stratagene). Each of the 95 PCR reactions consisted of 2 µl 10X KlenTaq PCR reaction

buffer (Clontech), 1.6 μ l dNTPs mix (2.5 mM each, Perkin-Elmer, Foster City, CA), 1 μ l sense primer, 15,398 (SEQ ID NO:15), 1 μ l antisense primer, ZC 15,397 (SEQ ID NO:16), 2 μ l RediLoad (Research Genetics, Inc.), 0.4 μ l 50X Advantage

5 KlenTaq Polymerase Mix (Clontech), 25 ng of DNA from an individual hybrid clone or control and ddH₂O for a total volume of 20 μ l. The reactions were overlaid with an equal amount of mineral oil and sealed. The PCR cycler conditions were as follows: an initial 1 cycle 5 minute 10 denaturation at 95°C, 35 cycles of a 1 minute denaturation at 95°C, 1 minute annealing at 62°C and 1.5 minute extension at 72°C, followed by a final 1 cycle extension of 7 minutes at 72°C. The reactions were separated by electrophoresis on a 2% agarose gel (Life Technologies, Gaithersburg, MD).

15 The results showed that ZSMF-3 maps 384.88 cR_3000 from the top of the human chromosome 5 linkage group on the WICGR radiation hybrid map. The proximal and distal framework markers were D5S492 and WI-3394, respectively. The use of surrounding markers positions 20 ZSMF-3 in the 5q23.1 - q23.2 region on the integrated LDB chromosome 5 map (The Genetic Location Database, University of Southampton, WWW server:http://cedar.genetics.soton.ac.uk/public_html/).

Example 4
ZSMF-3 Adenovirus Expression Vector Construction

5 Production of adenovirus containing ZSMF-3 NF and CF was done according to the procedure of He et al., Proc. Natl. Acad. Sci. USA 95: 2509-14, 1998. Briefly, the cDNA encoding zsmf3 was amplified by PCR using primers ZC 17960 (SEQ ID NO:38) and ZC 17961 (SEQ ID NO:39). The oligos added two unique restriction sites, FseI and AscI. The reaction proceeded a 95°C for 5 minutes, followed by 25 cycles at 95°C for 1 minute, 58°C for 1 minute and 72°C for 1.5 minutes, with a 7 minute extension at 72°C. The PCR product was visualized by gel electrophoresis, excised and purified. The zsmf3 cDNA was digested with the restriction enzymes FseI and AscI, phenol/chloroform extracted, ethanol precipitated, and rehydrated in 20 µl TE (10 mM Tris HCl/1 mM EDTA pH 8.0). The zsmf3 fragment was then ligated into the FseI-AscI sites of the transgenic vector pTG12-8 and transformed into DH10B competent cells by electroporation. Clones containing zsmf3 were identified by mini prep followed by digestion with FseI-AscI. Identity of the clone was confirmed by sequence analysis.

25 Preparation of DNA construct for generation of Adenovirus

A 966 base pair ZSMF-3 fragment (nucleotide 1-966 of SEQ ID NO:1) was released from the TG12-8 vector using FseI and AscI restriction enzymes. The cDNA was isolated by electrophoresis on a 1% low melt SeaPlaque GTG gel. A band of the appropriate size was excised from the gel and the DNA extracted.

The zsmf3 cDNA was cloned into the FseI-AscI sites of pAdTrack CMV vector (He, et al., ibid.) in which the native polylinker was replaced with FseI, EcoRV, AscI sites. Ligation was performed using the Fast-Link DNA ligation and screening kit by Epicentre Technologies (Madison, WI). Five micrograms of the pAdTrack CMV zsmf3

plasmid was digested with PmeI to linearize the plasmid. One microgram of the linearized plasmid was co-transformed with 200 ng of supercoiled pAdEasy (He et al., *ibid.*) into BJ5183 cells (He et al., *ibid.*) by electroporation and 5 plated onto plates coated with LB containing 25 µg/ml kanamycin. The smallest colonies were picked and expanded in LB/kanamycin and recombinant adenovirus DNA identified by standard mini prep procedures. Digestion of the recombinant adenovirus DNA with FseI-AsCI confirmed the 10 presence of ZSMF-3. The recombinant adenovirus mini prep DNA was transformed into DH10B competent cells and DNA prepared using a Plasmid Maxi Kit (Qiagen) according to manufacturer's instructions.

15 Transfection of 293a Cells with Recombinant DNA

Five micrograms of recombinant adenoviral DNA was digested with PacI enzyme (NEBiolabs, Beverly, MA) for 3 hours at 37°C in a reaction volume of 100 µl containing 20-20 30U of PacI. The digested DNA was extracted twice with an equal volume of phenol/chloroform and precipitate with ethanol. The DNA pellet was resuspended in 10 µl distilled water. A T25 flask of QBI-293A cells (Quantum Biotechnologies, Inc. Montreal, QC. Canada) set up the 25 day before which were 60-70% confluent were transfected with the PacI digested DNA. The PacI-digested DNA was diluted up to a total volume of 50 µl with sterile HBS (150 mM NaCl, 20 mM HEPES). In a separate tube, 20 µl DOTAP (Boehringer Mannheim, 1 µg/µl) was diluted to a total 30 volume of 100 µl with HBS. The DNA was added to the DOTAP, mixed gently by pipetting up and down, and left at room temperature for 15 minutes. The media was removed from the 293A cells and washed with 5 mls serum-free MEMalpha (GIBCO BRL, Gaithersburg, MD) containing 1 mM Sodium Pyruvate 35 (GIBCO BRL), 0.1 mM MEM non-essential amino acids (GIBCO BRL) and 25 mM HEPES buffer (GIBCO BRL). Five milliliters

of serum-free MEM was added to the 293A cells and held at 37°C. The DNA/lipid mixture was added drop-wise to the T25 flask of 293A cells, mixed gently and incubated at 37°C for 4 hours. After 4 hours the media containing the 5 DNA/lipid mixture was aspirated off and replaced with 5 mls complete MEM containing 5% fetal bovine serum. The transfected cells were monitored for Green Fluorescent Protein (GFP) expression and formation of viral plaques.

10 Crude rAdV Lysate

Seven days after transfection of 293A cells with the recombinant adenoviral DNA, the cells expressed the GFP protein and started to form foci. These foci are viral 15 "plaques" and the crude viral lysate was collected by using a cell scraper to collect all of the 293A cells. The lysate was transferred to a 50 ml conical tube. To release most of the virus particles from the cells, three 20 freeze/thaw cycles were done in a dry ice/ethanol bath and a 37° waterbath.

Primary (1°) Amplification of rAdV

The crude lysate was amplified to obtain a 25 working "stock" of rAdV lysate. Ten, 10 cm plates of nearly confluent (80-90%) 293A cells were set up 20 hours previously, 200 µl of crude rAdV lysate added to each 10 cm plate and monitored for 48 to 72 hours looking for CPE (Cytopathic Effect) under the white light microscope and 30 expression of GFP under the fluorescent microscope. When all of the 293A cells showed CPE this 1° stock lysate was collected and freeze/thaw cycles performed as described under Crude rAdV Lysate.

35 Secondary (2°) Amplification of rAdV

Twenty, 15 cm tissue culture dishes of 293A cells were prepared so that the cells were 80-90% confluent. All

but 20 mls of 5% MEM media was removed and each dish was inoculated with 300-500 μ l 1 $^{\circ}$ amplified rAdv lysate. After 48 hours the 293A cells were lysed from virus production and this lysate was collected into 250 ml polypropylene 5 centrifuge bottles and the rAdv purified.

Adv/cDNA Purification

NP-40 detergent was added to a final 10 concentration of 0.5% to the bottles of crude lysate in order to lyse all cells. Bottles were placed on a rotating platform for 10 minutes agitating as fast as possible without the bottles falling over. The debris was pelleted by centrifugation at 20,000 X G for 15 minutes. The 15 supernatant was transferred to 250 ml polycarbonate centrifuge bottles and 0.5 volumes of 20% PEG8000/2.5M NaCl solution added. The supernatant containing bottles were shaken overnight on ice. The bottles were centrifuged at 20,000 X G for 15 minutes and supernatant discarded. The 20 white precipitate (precipitated virus/PEG) was collected using a sterile cell scraper and the precipitate from 2 bottles was resuspended in 2.5 mls PBS. The virus solution was placed in 2 ml microcentrifuge tubes and centrifuged at 14,000 X G in the microfuge for 10 minutes to remove any 25 additional cell debris. The supernatant from the 2 ml microcentrifuge tubes was transferred into a 15 ml polypropylene snapcap tube and adjusted to a density of 1.34g/ml with Cesium Chloride (CsCl). The solution was transferred polycarbonate thick-walled centrifuge tubes 3.2 30 ml (Beckman Instruments, Palo Alto, CA) and spun at 80,000 rpm (348,000 X G) for 3-4 hours at 25 $^{\circ}$ C in a Beckman Optima TLX micro ultracentrifuge with a TLA-100.4 rotor. The virus formed a white band. Using wide bore pipette tips virus band was collected.

35

Desalting the Virus

Pharmacia PD-10 columns prepacked with Sephadex G-25M (Pharmacia) were used to desalt the virus preparation. The column was equilibrated with 20 mls of PBS. The virus was loaded and allow it to run into the column. Five milliliters of PBS was added to the column and fractions of 8-10 drop collected. The optical densities of 1:50 dilutions of each fraction was determined at 260 nm on a spectrophotometer. Those fractions having virus were pooled and the optical density of a 1:50 dilution determined.

TCID 50 Viral Titration Assay

Recombinant virus infectivity was measured according to the method of Quantum Biotechnologies, Inc. Briefly, two 96-well tissue culture plates were seeded with 1 X 10^4 293A cells per well in MEM containing 2% fetal bovine serum for each recombinant virus to be assayed. After 24 hours 10-fold dilutions of each virus from 1 X 10^{-2} to 1 X 10^{-14} were made in MEM containing 2% fetal bovine serum. One hundred microliters of each dilution was placed in each of 20 wells. After 5 days at 37°C, wells were read either positive or negative for Cytopathic Effect (CPE) and a value for "Plaque Forming Units/ml" was calculated

The zsmf3 adenovirus had a titer of 8.0 X 10^{11} pfu/ml.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is claimed is:

1. An isolated polypeptide comprising a sequence of amino acids that is 96% identical to the amino acid residues as shown in SEQ ID NO:2 from Phe, residue 20, through Arg, residue 321, wherein said polypeptide is a semaphorin.
2. An isolated polypeptide according to claim 1, wherein any difference between said amino acid sequence of said isolated polypeptide and said corresponding amino acid sequence of SEQ ID NO:2 is due to a conservative amino acid substitution.
3. An isolated polypeptide according to claim 1 wherein said polypeptide comprises amino acid residues 20 through 321 of SEQ ID NO:2.
4. An isolated polypeptide according to claim 1, further comprising a heterologous affinity tag or binding domain.
5. An isolated polypeptide according to claim 4, wherein said affinity tag is a Glu-Glu tag.
6. An isolated polypeptide having the amino acid sequence of SEQ ID NO:2.
7. An isolated polypeptide comprising the amino acid sequence of amino acid residue 1 to amino acid residue 19 of SEQ ID NO:2.
8. An isolated polypeptide consisting of a semaphorin domain, a transmembrane domain and a cytoplasmic domain, wherein said semaphorin domain comprises the amino acid sequence of SEQ ID NO:2.

9. A fusion protein consisting essentially of a first portion and a second portion joined by a peptide bond, said first portion comprising a polypeptide wherein said polypeptide comprises a sequence of amino acids that is 96% identical to the amino acid residues as shown in SEQ ID NO:2 from Phe, residue 20, through Arg, residue 321; wherein said polypeptide is a semaphorin; and

 said second portion comprising another polypeptide.

10. A fusion protein according to claim 9, wherein said second polypeptide is a transmembrane domain of a semaphorin protein.

11. A fusion protein according to claim 10, wherein said second polypeptide further comprises a cytoplasmic domain of a semaphorin protein.

12. A fusion protein comprising a secretory signal sequence having the amino acid sequence of amino acid residues 1 to 19 of SEQ ID NO:2, wherein said secretory signal sequence is operably linked to an additional polypeptide.

13. An isolated polynucleotide selected from the group consisting of:

 a) a polynucleotide encoding a polypeptide comprising a sequence of amino acids that is 96% identical to the amino acid residues as shown in SEQ ID NO:2 from Phe, residue 20, through Arg, residue 321;

 b) degenerate polynucleotide sequences of a); and

 c) a polynucleotide sequence complementary to a) or

b);

 wherein said polynucleotide encodes a polypeptide which is a semaphorin.

14. An isolated polynucleotide according to claim 13 wherein said polypeptide comprises residues 20 through 321 of SEQ ID NO:2.

15. An isolated polynucleotide according to claim 13 wherein said protein comprises residues 1 through 321 of SEQ ID NO:2.

16. An expression vector comprising the following operably linked elements:

a transcription promoter;

a DNA segment encoding a polypeptide according to claim 1; and

a transcription terminator.

17. An expression vector according to claim 16 further comprising a secretory signal sequence operably linked to said DNA segment.

18. An expression vector according to claim 16 further comprising an affinity tag operably linked to said DNA segment.

19. A cultured cell containing an expression vector according to claim 16 wherein said cell expresses said DNA segment.

20. A cultured cell according to claim 19 wherein the expression vector further comprises a secretory signal sequence operably linked to said DNA segment and the cell secretes said polypeptide.

21. A cultured cell according to claim 19 wherein the expression vector further comprises an affinity tag operably linked to said DNA segment.

22. A method of making a semaphorin protein comprising:

- (a) culturing a host cell containing an expression vector according to claim 16; and
- (b) recovering said protein encoded by said DNA segment.

23. A method according to claim 22 wherein the expression vector further comprises a secretory signal sequence operably linked to said DNA segment, the cell secretes the protein into a culture medium, and the protein is recovered from the medium.

24. A method according to claim 22 wherein the expression vector further comprises an affinity tag operably linked to said DNA segment.

25. An antibody or antibody fragment that specifically binds to a polypeptide according to claim 1.

26. An antibody according to claim 25, wherein said antibody is selected from the group consisting of:

- a) polyclonal antibody;
- b) murine monoclonal antibody;
- c) humanized antibody derived from b); and
- d) human monoclonal antibody.

27. An antibody fragment according to claim 25, wherein said antibody fragment is selected from the group consisting of F(ab'), F(ab), Fab', Fab, Fv, scFv, and minimal recognition unit.

28. An anti-idiotype antibody that specifically binds to said antibody of claim 25.

29. A binding protein that specifically binds to a polypeptide according to claim 1.

30. A pharmaceutical composition comprising a polypeptide according to Claim 1, in combination with a pharmaceutically acceptable vehicle.

1 / 3

1	15 16	30 31	45 46	60 61	75 76
Murine SEM G	-	-MVVP GPLALSLLSSL TTL	-VSHLSSSQD	IASESSSEQMC TRR	EHP
Murine SEM F	-	-MKGA CILAWLFFSSLGVWRL	ARP	-	ETQDPAKQRA EHP
Murine ZSMF -3	AAACPSMHRKG.	SRE ASPPAEP T	-MRPA ALLLC LTLHAGA	-	GFPEDSEPI SIGHN YTK
Human ZSMF -3	-	-EFG LSPPAEP T	-MRSE ALLLYFTLLHFAGA	-	GFPEDSEPI SIGHN YTK
					*
		+ peptide	- ->		
		+ Starting Met	and signal	sequence	
90 91	105 106	120 121	135 136	150 151	
Murine SEM G	IVAFEDLKPWVF	NFTYPGV RQDFSQL AL	DPSRNQL I VGG RNYL	FRLSL	-ANVSL
Murine SEM F	VVSYKEIGPWLR	E FRAENAVDF SRL TF	DPGQKEL VVGARNYL	FRREL	-EDLSL
Murine ZSMF -3	QYPVFVGHKPGR	NTTQRHRLDIQMIMI	MNRT	-LYVAAARDHI	YTVDIDTSHTEE IYC
Human ZSMF -3	QYPVFVGHKPGR	NTTQRHRLDIQMIMI	MNGT	-LYVAAARDHI	YTVDIDTSHTEE IYC
165 166	180 181	195 196	210 211	225 226	
Murine SEM G	KTEEECQNYV	RVLIV-SGRKVF MCG	TNAFSPV CSSRQVGN	LSRTIEKINGVARCP	YDPRHN STAVI SSGQ
Murine SEM F	KSKEECQNYI	RVLLV-GGDRLLFTCG	TNAFTP VCT1RSLSN	LTE1HDQ1SGGMARCP	YSPQHN STALLTASG
Murine ZSMF -3	KHKDECHNF1	KVLLKKNDDTLFVCG	TNAFNPSCRN YVRDT	LETFGDEFSGGMARCP	YDAKHANIA LF-ADG
Human ZSMF -3	KHKDECHNF1	KVLLKKNDDA LFVCG	TNAFNPSCRN YKMDT	LEPF GDEFSGGMARCP	YDAKHANIA LF-ADG
					*

Fig. 1a

Murine SEM G	DFSGRDP	AIYRSLGSGPPLRTA	QYN SKWLNEPNFVAA	FDIGLFAYFFLRENA	VE -HDCGRTVYSRVA	RVCKNDVGG-R
Murine SEM F	DFPGRDP	AIYRSLGTLPPPLRTA	QYN SKWLNEPNFVSS	YDIGNFTYFFFRENA	VE -HDCGKTVFSRPA	RVCKNDIGG-R
Murine ZSMF-3	DFAIDA	VIYRSLGDSPTLRTV	KHDSKWLKEPYFVQA	VDYGDYIYFFFREIA	VEYNTMGKVVFPRVA	QVCKNDMGGSQ
Human ZSMF-3	DFAIDA	VIYRSLGEHSLRTV	KHDSKWLKEPYFVQA	VDYGDYIYFFFREIA	VEYNTMGKVVFPRVA	QVCKNDMGGSQ
		*				
Murine SEM G	315 316	330 331	345 346	360	361	375 376
Murine SEM F	FLLE	DTWTTFMKARLNC SR	PGEVPFYYNELQSAF	HLPEQ---DLIYGVF	TTNVNSIAASAVCAF	NLSAISKA FNGPFR
Murine ZSMF-3	FLLE	DTWTTFMKARLNC SR	PGEVPFYYNELQGT F	FLPEL---DLIYGF	TTNVNSIAASAVCF	NLSAISQAFNGDFK
Human ZSMF-3	RVLE	KQWTSFLKARLNC SV	PGDSHFYFNILQAVT	DVIRINGRDVVLATF	STPYNSPRAEF	
			PGDSHFYFNILQAVT	DVIRINGRDVVLATF	STPYNR .SCPSCVDL	INSSWLPSPRGFSS
		*		*		
Murine SEM G	390	391	405	406		
Murine SEM F	Y	Y	QENPRAAWLPIANPI			
Murine ZSMF-3			QENSRSAMLPYNPN			
Human ZSMF-3	N	KPYLQLKKKKKRAA				

+ Stop

Fig. 1b

3 / 3

ZSMF -3	EFGLSPPAEPMTMRSEALLLYFTLLHFAAGFPEDSEPI SISHGNYTQYPVFGHKPGRNTTQRHRLDIQMIMMN	10 20 30 40 50 60 70
SEMA IVa	AAAACPSMHRKGSSREASPPAEPMTMRPAALLLCLTLHCAGAGFPEDSEPI SISHGNYTQYPVFGHKPGRNTTQRHRLDIQMIMMN	10 20 30 40 50 60 70 80
ZSMF -3	GTYIARDHIYTVIDTSHTEEYCSKKLTWKSROADVDTCRMKGKHKDECHNFIKVLLKNDTAFVCGTNAFNPSCRNYKMDTLEP	80 90 100 110 120 130 140 150 160
SEMA IVa	RTLYVARDHIYTVIDTSHTEEYCSKKLTWKSROADVDTCRMKGKHKDECHNFIKVLLKNDTFLVCGTNAFNPSCRNYRVDTLET	90 100 110 120 130 140 150 160 170
ZSMF -3	FGDEFGSMARCPYDAKHANVALFADGKLYSATVTDLAIADAVIYRSLGEHSHSLRTVKHDSKWLKEPYFVQAVDYGDYIYFFFREIAVEY	170 180 190 200 210 220 230 240 250
SEMA IVa	FGDEFGSMARCPYDAKHANIALFADGKLYSATVTDLAIADAVIYRSLGDSPLRTVKHDSKWLKEPYFVQAVDYGDYIYFFFREIAVEY	180 190 200 210 220 230 240 250 260
ZSMF -3	NTMGKVVFPRTVAQVCKNDMGGSQRVLKEQWTSFLKARLNCSVPGDSHFYFNILQAVTDVIRINGRDVVLATFSTPYNSIPGSAVCAYDML	260 270 280 290 300 310 320 330 340 350
SEMA IVa	NTMGKVVFPRTVAQVCKNDMGGSQRVLKEQWTSFLKARLNCSVPGDSHFYFNILQAVTDVIRINGRDVVLATFSTPYNSIPGSAVCAYDML	270 280 290 300 310 320 330 340 350
ZSMF -3	WLPSPRGFSSNKPYQLKKKKRAA	350 360
SEMA IVa	DIANVFTGRFKEQKSPDSTWTPVPD	360 370 380

Fig. 2

SEQUENCE LISTING

<110> ZymoGenetics, Inc.
1201 Eastlake Avenue East
Seattle, Washington 98102
United States of America

<120> SEMAPHORIN ZSMF-3

<130> 97-58

<150> 60/069,871
<151> 1997-12-19

<160> 39

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 1054
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)...(963)

<400> 1

atg agg tca gaa gcc ttg ctg cta tat ttc aca ctg cta cac ttt gct 48
Met Arg Ser Glu Ala Leu Leu Leu Tyr Phe Thr Leu Leu His Phe Ala
1 5 10 15

ggg gct ggt ttc cca gaa gat tct gag cca atc agt att tcg cat ggc 96
Gly Ala Gly Phe Pro Glu Asp Ser Glu Pro Ile Ser Ile Ser His Gly
20 25 30

aac tat aca aaa cag tat ccg gtg ttt gtg ggc cac aag cca gga cgg 144
Asn Tyr Thr Lys Gln Tyr Pro Val Phe Val Gly His Lys Pro Gly Arg
35 40 45

aac acc aca cag agg cac agg ctg gac atc cag atg att atg atc atg 192

Asn Thr Thr Gln Arg His Arg Leu Asp Ile Gln Met Ile Met Ile Met
50 55 60

aac gga acc ctc tac att gct gct agg gac cat att tat act gtt gat 240
Asn Gly Thr Leu Tyr Ile Ala Ala Arg Asp His Ile Tyr Thr Val Asp
65 70 75 80

ata gac aca tca cac acg gaa gaa att tat tgt agc aaa aaa ctg aca 288
Ile Asp Thr Ser His Thr Glu Glu Ile Tyr Cys Ser Lys Lys Leu Thr
85 90 95

tgg aaa tct aga cag gcc gat gta gac aca tgc aga atg aag gga aaa 336
Trp Lys Ser Arg Gln Ala Asp Val Asp Thr Cys Arg Met Lys Gly Lys
100 105 110

cat aag gat gag tgc cac aac ttt att aaa gtt ctt cta aag aaa aac 384
His Lys Asp Glu Cys His Asn Phe Ile Lys Val Leu Leu Lys Lys Asn
115 120 125

gat gat gca ttg ttt gtc tgt gga act aat gcc ttc aac cct tcc tgc 432
Asp Asp Ala Leu Phe Val Cys Gly Thr Asn Ala Phe Asn Pro Ser Cys
130 135 140

aga aac tat aag atg gat aca ttg gaa cca ttc ggg gat gaa ttc agc 480
Arg Asn Tyr Lys Met Asp Thr Leu Glu Pro Phe Gly Asp Glu Phe Ser
145 150 155 160

gga atg gcc aga tgc cca tat gat gcc aaa cat gcc aac gtt gca ctg 528
Gly Met Ala Arg Cys Pro Tyr Asp Ala Lys His Ala Asn Val Ala Leu
165 170 175

ttt gca gat gga aaa cta tac tca gcc aca gtg act gac ttc ctt gcc 576
Phe Ala Asp Gly Lys Leu Tyr Ser Ala Thr Val Thr Asp Phe Leu Ala
180 185 190

att gac gca gtc att tac cgg agt ctt gga gaa agc cct acc ctg cgg 624
Ile Asp Ala Val Ile Tyr Arg Ser Leu Gly Glu Ser Pro Thr Leu Arg
195 200 205

acc gtc aag cac gat tca aaa tgg ttg aaa gaa cca tac ttt gtt caa 672
Thr Val Lys His Asp Ser Lys Trp Leu Lys Glu Pro Tyr Phe Val Gln
210 215 220

gcc gtg gat tac gga gat tat atc tac ttc ttc agg gaa ata gca 720
Ala Val Asp Tyr Gly Asp Tyr Ile Tyr Phe Phe Arg Glu Ile Ala

225	230	235	240													
gtg gag tat aac acc atg gga aag gta gtt ttc cca aga gtg gct cag				768												
Val	Glu	Tyr	Asn	Thr	Met	Gly	Lys	Val	Val	Phe	Pro	Arg	Val	Ala	Gln	
245				250	255											
gtt tgt aag aat gat atg gga gga tct caa aga gtc ctg gag aaa cag				816												
Val	Cys	Lys	Asn	Asp	Met	Gly	Gly	Ser	Gln	Arg	Val	Leu	Glu	Lys	Gln	
260				265	270											
tgg acg tcg ttc ctg aag gcg cgc ttg aac tgc tca gtt cct gga gac				864												
Trp	Thr	Ser	Phe	Leu	Lys	Ala	Arg	Leu	Asn	Cys	Ser	Val	Pro	Gly	Asp	
275				280	285											
tct cat ttt tat ttc aac att ctc cag gca gtt aca gat gtg att cgt				912												
Ser	His	Phe	Tyr	Phe	Asn	Ile	Leu	Gln	Ala	Val	Thr	Asp	Val	Ile	Arg	
290				295	300											
atc aac ggg cgt gat gtt gtc ctg gca acg ttt tct aca cct tat aac				960												
Ile	Asn	Gly	Arg	Asp	Val	Val	Leu	Ala	Thr	Ser	Thr	Pro	Tyr	Asn		
305				310	315	320										
agg taatcatgcc ctagctgtgt tgacctcatc aattcttcct ggtttccttc				1013												
Arg																
tcccaggggg ttctttcta ataaaccata tttgcaactg a				1054												
<210> 2																
<211> 321																
<212> PRT																
<213> Homo sapiens																
<400> 2																
Met	Arg	Ser	Glu	Ala	Leu	Leu	Tyr	Phe	Thr	Leu	Leu	His	Phe	Ala		
1				5				10				15				
Gly	Ala	Gly	Phe	Pro	Glu	Asp	Ser	Glu	Pro	Ile	Ser	Ile	Ser	His	Gly	
				20				25				30				
Asn	Tyr	Thr	Lys	Gln	Tyr	Pro	Val	Phe	Val	Gly	His	Lys	Pro	Gly	Arg	
				35				40			45					
Asn	Thr	Thr	Gln	Arg	His	Arg	Leu	Asp	Ile	Gln	Met	Ile	Met	Ile	Met	
				50				55			60					
Asn	Gly	Thr	Leu	Tyr	Ile	Ala	Ala	Arg	Asp	His	Ile	Tyr	Thr	Val	Asp	
				65				70			75			80		
Ile	Asp	Thr	Ser	His	Thr	Glu	Glu	Ile	Tyr	Cys	Ser	Lys	Lys	Leu	Thr	

85	90	95
Trp Lys Ser Arg Gln Ala Asp Val Asp Thr Cys Arg Met Lys Gly Lys		
100	105	110
His Lys Asp Glu Cys His Asn Phe Ile Lys Val Leu Leu Lys Lys Asn		
115	120	125
Asp Asp Ala Leu Phe Val Cys Gly Thr Asn Ala Phe Asn Pro Ser Cys		
130	135	140
Arg Asn Tyr Lys Met Asp Thr Leu Glu Pro Phe Gly Asp Glu Phe Ser		
145	150	155
160		
Gly Met Ala Arg Cys Pro Tyr Asp Ala Lys His Ala Asn Val Ala Leu		
165	170	175
Phe Ala Asp Gly Lys Leu Tyr Ser Ala Thr Val Thr Asp Phe Leu Ala		
180	185	190
Ile Asp Ala Val Ile Tyr Arg Ser Leu Gly Glu Ser Pro Thr Leu Arg		
195	200	205
Thr Val Lys His Asp Ser Lys Trp Leu Lys Glu Pro Tyr Phe Val Gln		
210	215	220
Ala Val Asp Tyr Gly Asp Tyr Ile Tyr Phe Phe Arg Glu Ile Ala		
225	230	235
240		
Val Glu Tyr Asn Thr Met Gly Lys Val Val Phe Pro Arg Val Ala Gln		
245	250	255
Val Cys Lys Asn Asp Met Gly Gly Ser Gln Arg Val Leu Glu Lys Gln		
260	265	270
Trp Thr Ser Phe Leu Lys Ala Arg Leu Asn Cys Ser Val Pro Gly Asp		
275	280	285
Ser His Phe Tyr Phe Asn Ile Leu Gln Ala Val Thr Asp Val Ile Arg		
290	295	300
Ile Asn Gly Arg Asp Val Val Leu Ala Thr Phe Ser Thr Pro Tyr Asn		
305	310	315
Arg		

<210> 3

<211> 1047

<212> DNA

<213> Mus musculus

<220>

<221> CDS

<222> (70)...(1032)

<400> 3

gcggccgcct gccccgtcgat gcaccgaaaa gggtaatgtt gagaagcaag tccccccgt
 gAACCTACT atg cgg cca gca gcc tta ctg ctg tgt ctc aca ctg cta cac
 Met Arg Pro Ala Ala Leu Leu Leu Cys Leu Thr Leu Leu His

60

111

1	5	10	
tgc gct ggg gct ggt ttc cca gaa gat tcc gag cca atc agt att tcg Cys Ala Gly Ala Gly Phe Pro Glu Asp Ser Glu Pro Ile Ser Ile Ser			159
15	20	25	30
cat ggc aac tat aca aaa cag tat ccg gtg ttt gtg ggc cac aag cca His Gly Asn Tyr Thr Lys Gln Tyr Pro Val Phe Val Gly His Lys Pro			207
35	40	45	
gga cgg aac acc acg cag agg cac agg ctg gac atc cag atg atc atg Gly Arg Asn Thr Thr Gln Arg His Arg Leu Asp Ile Gln Met Ile Met			255
50	55	60	
atc atg aac aga acc ctc tac gtt gct gct cga gac cat att tat act Ile Met Asn Arg Thr Leu Tyr Val Ala Ala Arg Asp His Ile Tyr Thr			303
65	70	75	
gtt gat ata gac aca tcc cac aca gaa gaa att tac tgt agc aaa aaa Val Asp Ile Asp Thr Ser His Thr Glu Glu Ile Tyr Cys Ser Lys Lys			351
80	85	90	
ctg aca tgg aaa tct aga cag gct gac gta gac aca tgc agg atg aag Leu Thr Trp Lys Ser Arg Gln Ala Asp Val Asp Thr Cys Arg Met Lys			399
95	100	105	110
ggg aaa cat aag gat gaa tgt cac aac ttc att aaa gtt ctt ctc aag Gly Lys His Lys Asp Glu Cys His Asn Phe Ile Lys Val Leu Leu Lys			447
115	120	125	
aag aat gat gat acg ctg ttt gtc tgt gga acc aat gcc ttc aac cct Lys Asn Asp Asp Thr Leu Phe Val Cys Gly Thr Asn Ala Phe Asn Pro			495
130	135	140	
tcc tgc aga aac tac agg gtc gat acc ttg gaa act ttt ggg gat gaa Ser Cys Arg Asn Tyr Arg Val Asp Thr Leu Glu Thr Phe Gly Asp Glu			543
145	150	155	
ttt agc gga atg gcc aga tgc cct tat gat gcc aaa cat gcc aac atc Phe Ser Gly Met Ala Arg Cys Pro Tyr Asp Ala Lys His Ala Asn Ile			591
160	165	170	
gct ctg ttt gca gat gga aaa ctc tac tcg gct aca gtg act gac ttt Ala Leu Phe Ala Asp Gly Lys Leu Tyr Ser Ala Thr Val Thr Asp Phe			639
175	180	185	190

ctg gcc att gat gca gtc att tac agg agc ctc gga gac agc cct acc Leu Ala Ile Asp Ala Val Ile Tyr Arg Ser Leu Gly Asp Ser Pro Thr	195	200	205	687	
ctc agg act gtc aag cat gat tca aag tgg ttg aaa gag ccg tac ttt Leu Arg Thr Val Lys His Asp Ser Lys Trp Leu Lys Glu Pro Tyr Phe	210	215	220	735	
gtc caa gcc gtg gat tat ggg gac tat atc tac ttc ttc ttc aga gaa Val Gln Ala Val Asp Tyr Gly Asp Tyr Ile Tyr Phe Phe Arg Glu	225	230	235	783	
att gca gta gaa tac aac act atg ggg aag gtt gtt ttc cct agg gtg Ile Ala Val Glu Tyr Asn Thr Met Gly Lys Val Val Phe Pro Arg Val	240	245	250	831	
gct cag gtc tgt aag aat gac atg gga ggg tct cag aga gtc ctg gag Ala Gln Val Cys Lys Asn Asp Met Gly Gly Ser Gln Arg Val Leu Glu	255	260	265	270	879
aag cag tgg aca tct ttc ctg aag gct cgc ctg aac tgc tcg gtg cct Lys Gln Trp Thr Ser Phe Leu Lys Ala Arg Leu Asn Cys Ser Val Pro	275	280	285	927	
gga gac tct cat ttt tat ttc aat ata ctc cag gca gtt aca gat gtg Gly Asp Ser His Phe Tyr Phe Asn Ile Leu Gln Ala Val Thr Asp Val	290	295	300	975	
att cgc att aat ggc cgt gat gtt gtc ttg gca acc ttt tcc aca cct Ile Arg Ile Asn Gly Arg Asp Val Val Leu Ala Thr Phe Ser Thr Pro	305	310	315	1023	
tat aac agc cctcgtgccg aattc Tyr Asn Ser	320			1047	

<210> 4
<211> 321
<212> PRT
<213> Mus musculus

<400> 4
Met Arg Pro Ala Ala Leu Leu Leu Cys Leu Thr Leu Leu His Cys Ala

1	5	10	15												
Gly	Ala	Gly	Phe	Pro	Glu	Asp	Ser	Glu	Pro	Ile	Ser	Ile	Ser	His	Gly
			20			25								30	
Asn	Tyr	Thr	Lys	Gln	Tyr	Pro	Val	Phe	Val	Gly	His	Lys	Pro	Gly	Arg
			35			40								45	
Asn	Thr	Thr	Gln	Arg	His	Arg	Leu	Asp	Ile	Gln	Met	Ile	Met	Ile	Met
			50			55					60				
Asn	Arg	Thr	Leu	Tyr	Val	Ala	Ala	Arg	Asp	His	Ile	Tyr	Thr	Val	Asp
	65				70					75				80	
Ile	Asp	Thr	Ser	His	Thr	Glu	Glu	Ile	Tyr	Cys	Ser	Lys	Lys	Leu	Thr
			85			90								95	
Trp	Lys	Ser	Arg	Gln	Ala	Asp	Val	Asp	Thr	Cys	Arg	Met	Lys	Gly	Lys
			100			105						110			
His	Lys	Asp	Glu	Cys	His	Asn	Phe	Ile	Lys	Val	Leu	Leu	Lys	Lys	Asn
	115				120							125			
Asp	Asp	Thr	Leu	Phe	Val	Cys	Gly	Thr	Asn	Ala	Phe	Asn	Pro	Ser	Cys
			130			135					140				
Arg	Asn	Tyr	Arg	Val	Asp	Thr	Leu	Glu	Thr	Phe	Gly	Asp	Glu	Phe	Ser
	145				150					155				160	
Gly	Met	Ala	Arg	Cys	Pro	Tyr	Asp	Ala	Lys	His	Ala	Asn	Ile	Ala	Leu
					165				170				175		
Phe	Ala	Asp	Gly	Lys	Leu	Tyr	Ser	Ala	Thr	Val	Thr	Asp	Phe	Leu	Ala
			180			185						190			
Ile	Asp	Ala	Val	Ile	Tyr	Arg	Ser	Leu	Gly	Asp	Ser	Pro	Thr	Leu	Arg
			195			200						205			
Thr	Val	Lys	His	Asp	Ser	Lys	Trp	Leu	Lys	Glu	Pro	Tyr	Phe	Val	Gln
			210			215				220					
Ala	Val	Asp	Tyr	Gly	Asp	Tyr	Ile	Tyr	Phe	Phe	Arg	Glu	Ile	Ala	
	225				230				235				240		
Val	Glu	Tyr	Asn	Thr	Met	Gly	Lys	Val	Val	Phe	Pro	Arg	Val	Ala	Gln
					245				250				255		
Val	Cys	Lys	Asn	Asp	Met	Gly	Gly	Ser	Gln	Arg	Val	Leu	Glu	Lys	Gln
					260			265				270			
Trp	Thr	Ser	Phe	Leu	Lys	Ala	Arg	Leu	Asn	Cys	Ser	Val	Pro	Gly	Asp
					275			280				285			
Ser	His	Phe	Tyr	Phe	Asn	Ile	Leu	Gln	Ala	Val	Thr	Asp	Val	Ile	Arg
	290				295						300				
Ile	Asn	Gly	Arg	Asp	Val	Val	Leu	Ala	Thr	Phe	Ser	Thr	Pro	Tyr	Asn
	305				310				315				320		
Ser															

<210> 5

<211> 963

<212> DNA

<213> Artificial Sequence

<220>

<223> Degenerate oligonucleotide sequence encoding ZSMF-3 polypeptide of SEQ ID NO:2.

<221> variation

<222> (1)...(963)

<223> Each N is independently any nucleotide.

<400> 5

atgmgnwsng argcnytnyt nytnaytty acnytnytnc ayttgcngg ngcnggntty	60
ccngargayw sngarccnat hwsnathwsn cayggnaayt ayacnaarca rtayccngt	120
ttygtngnc ayaarcnngg nmgnayacn acncarmgnc aymgnytng yathcaratg	180
athatgatha tgaayggnc nytnayath gcngcnmgng aycayathta yacngtngay	240
athgayacnw sncayacnra rgarathtay tgywsnaara arytnacntg gaarwsnmgn	300
cargcngayg tngayacntg ymgnatgaar gnaarcaya argaygartg ycayaayt	360
athaargtny tnytnaaraa raaygaygay gcnytnatty tntgyggnc naaygcntt	420
aayccnwsnt gymgnayta yaaratggay acnytngarc cnytggng ygarttywsn	480
gnatggcnm gntgycnta ygaygcnaar caygcnaayg tngcnytnnt ygcngayggn	540
aarytnayw sngcnacngt nacngaytty ytnngnathg aygcngtnat htaymgnwsn	600
ytnngngarw snccnacnyt nmgnacngt aarcaygayw snaartggyt naargarccn	660
taytgytnc argcngtng aytayggngay tayathtayt tytgytng ygarathgcn	720
gtngartaya ayacnatggg naargtngt tnytggng tngcncargt ntgyaaraay	780
gayatggng gnwsncarmg ngtnyngar aarcartgga cnwsntyyt naargcnmgn	840
ytnaaytgyw sngtnccnng ngaywsncay tnytaytta ayathytnca rgcngtnacn	900
gaygttnathm gnathaaygg nmgnaygtn gtnytnacn cnytwsnac nccntayaay	960
mgn	963

<210> 6

<211> 199

<212> DNA

<213> NorthArtificial Sequence

<220>

<223> Human ZSMF-3 probe for Northern Blot Analysis

<400> 6

cgagaatgcc ttcaaccctt cctgcagaaa ctataagatg gatacattgg aaccattcgg	60
ggatgaattc agcggaatgg ccagatgccc atatgatgcc aaacatgcc aacgttgcact	120
gtttgcagat ggaaaactat actcagccac agtgaactgac ttccttgcca ttgacgcagc	180
atttaccgga gtcttggag	199

<210> 7

<211> 20

<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC694

<400> 7
taatacgact cactataggg

20

<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC976

<400> 8
cgttgtaaaa cgacggcc

18

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC14201

<400> 9
tatagacaca tcccacacag

20

<210> 10
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC14202

<400> 10
cgaggctcct gtaaatgact

20

<210> 11
<211> 23
<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14095

<400> 11

ctccaagact ccgttaaatg ctg

23

<210> 12

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14096

<400> 12

cgagaatgcc ttcaaccctt cc

22

<210> 13

<211> 407

<212> PRT

<213> Mus musculus

<400> 13

Met	Val	Val	Pro	Gly	Pro	Leu	Ala	Leu	Ser	Leu	Leu	Leu	Ser	Ser	Leu
1						5			10				15		
Thr	Leu	Leu	Val	Ser	His	Leu	Ser	Ser	Ser	Gln	Asp	Ile	Ala	Ser	Glu
								20		25			30		
Ser	Ser	Ser	Glu	Gln	Gln	Met	Cys	Thr	Arg	Arg	Glu	His	Pro	Ile	Val
						35			40			45			
Ala	Phe	Glu	Asp	Leu	Lys	Pro	Trp	Val	Phe	Asn	Phe	Thr	Tyr	Pro	Gly
						50		55			60				
Val	Arg	Asp	Phe	Ser	Gln	Leu	Ala	Leu	Asp	Pro	Ser	Arg	Asn	Gln	Leu
65								70			75		80		
Ile	Val	Gly	Gly	Arg	Asn	Tyr	Leu	Phe	Arg	Leu	Ser	Leu	Ala	Asn	Val
						85			90			95			
Ser	Leu	Leu	Gln	Ala	Thr	Glu	Trp	Ala	Ser	Ser	Glu	Asp	Thr	Arg	Arg
						100			105			110			
Ser	Cys	Gln	Ser	Lys	Gly	Lys	Thr	Glu	Glu	Glu	Cys	Gln	Asn	Tyr	Val
						115			120			125			
Arg	Val	Leu	Ile	Val	Ser	Gly	Arg	Lys	Val	Phe	Met	Cys	Gly	Thr	Asn
						130		135			140				
Ala	Phe	Ser	Pro	Val	Cys	Ser	Ser	Arg	Gln	Val	Gly	Asn	Leu	Ser	Arg
145						150				155			160		

Thr Ile Glu Lys Ile Asn Gly Val Ala Arg Cys Pro Tyr Asp Pro Arg
 165 170 175
 His Asn Ser Thr Ala Val Ile Ser Ser Gln Gly Glu Leu Tyr Ala Ala
 180 185 190
 Thr Val Ile Asp Phe Ser Gly Arg Asp Pro Ala Ile Tyr Arg Ser Leu
 195 200 205
 Gly Ser Gly Pro Pro Leu Arg Thr Ala Gln Tyr Asn Ser Lys Trp Leu
 210 215 220
 Asn Glu Pro Asn Phe Val Ala Ala Phe Asp Ile Gly Leu Phe Ala Tyr
 225 230 235 240
 Phe Phe Leu Arg Glu Asn Ala Val Glu His Asp Cys Gly Arg Thr Val
 245 250 255
 Tyr Ser Arg Val Ala Arg Val Cys Lys Asn Asp Val Gly Gly Arg Phe
 260 265 270
 Leu Leu Glu Asp Thr Trp Thr Phe Met Lys Ala Arg Leu Asn Cys
 275 280 285
 Ser Arg Pro Gly Glu Val Pro Phe Tyr Tyr Asn Glu Leu Gln Ser Ala
 290 295 300
 Phe His Leu Pro Glu Gln Asp Leu Ile Tyr Gly Val Phe Thr Thr Asn
 305 310 315 320
 Val Asn Ser Ile Ala Ala Ser Ala Val Cys Ala Phe Asn Leu Ser Ala
 325 330 335
 Ile Ser Lys Ala Phe Asn Gly Pro Phe Arg Tyr Gln Glu Asn Pro Arg
 340 345 350
 Ala Ala Trp Leu Pro Ile Ala Asn Pro Ile Pro Asn Phe Gln Cys Gly
 355 360 365
 Thr Leu Pro Glu Thr Gly Pro Asn Glu Asn Leu Thr Glu Arg Ser Leu
 370 375 380
 Gln Asp Ala Gln Arg Leu Phe Leu Met Ser Glu Ala Val Gln Pro Val
 385 390 395 400
 Thr Pro Glu Pro Cys Val Thr
 405

<210> 14
 <211> 397
 <212> PRT
 <213> Mus musculus

<400> 14

Met	Lys	Gly	Ala	Cys	Ile	Leu	Ala	Trp	Leu	Phe	Ser	Ser	Leu	Gly	Val
1					5				10					15	
Trp	Arg	Leu	Ala	Arg	Pro	Glu	Thr	Gln	Asp	Pro	Ala	Lys	Cys	Gln	Arg
					20			25					30		
Ala	Glu	His	Pro	Val	Val	Ser	Tyr	Lys	Glu	Ile	Gly	Pro	Trp	Leu	Arg
					35			40					45		

Glu Phe Arg Ala Glu Asn Ala Val Asp Phe Ser Arg Leu Thr Phe Asp
50 55 60
Pro Gly Gln Lys Glu Leu Val Val Gly Ala Arg Asn Tyr Leu Phe Arg
65 70 75 80
Leu Glu Leu Glu Asp Leu Ser Leu Ile Gln Ala Val Glu Trp Glu Cys
85 90 95
Asp Glu Ala Thr Lys Lys Ala Cys Tyr Ser Lys Gly Lys Ser Lys Glu
100 105 110
Glu Cys Gln Asn Tyr Ile Arg Val Leu Leu Val Gly Gly Asp Arg Leu
115 120 125
Phe Thr Cys Gly Thr Asn Ala Phe Thr Pro Val Cys Thr Ile Arg Ser
130 135 140
Leu Ser Asn Leu Thr Glu Ile His Asp Gln Ile Ser Gly Met Ala Arg
145 150 155 160
Cys Pro Tyr Ser Pro Gln His Asn Ser Thr Ala Leu Leu Thr Ala Ser
165 170 175
Gly Glu Leu Tyr Ala Ala Thr Ala Met Asp Phe Pro Gly Arg Asp Pro
180 185 190
Ala Ile Tyr Arg Ser Leu Gly Thr Leu Pro Pro Leu Arg Thr Ala Gln
195 200 205
Tyr Asn Ser Lys Trp Leu Asn Glu Pro Asn Phe Val Ser Ser Tyr Asp
210 215 220
Ile Gly Asn Phe Thr Tyr Phe Phe Arg Glu Asn Ala Val Glu His
225 230 235 240
Asp Cys Gly Lys Thr Val Phe Ser Arg Pro Ala Arg Val Cys Lys Asn
245 250 255
Asp Ile Gly Gly Arg Phe Leu Leu Glu Asp Thr Trp Thr Phe Met
260 265 270
Lys Ala Arg Leu Asn Cys Ser Arg Pro Gly Glu Val Pro Phe Tyr Tyr
275 280 285
Asn Glu Leu Gln Gly Thr Phe Phe Leu Pro Glu Leu Asp Leu Ile Tyr
290 295 300
Gly Ile Phe Thr Thr Asn Val Asn Ser Ile Ala Ser Ser Ala Val Cys
305 310 315 320
Val Phe Asn Leu Ser Ala Ile Ser Gln Ala Phe Asn Gly Pro Phe Lys
325 330 335
Tyr Gln Glu Asn Ser Arg Ser Ala Trp Leu Pro Tyr Pro Asn Pro Asn
340 345 350
Pro Asn Phe Gln Cys Gly Thr Met Asp Gln Gly Leu Tyr Val Asn Leu
355 360 365
Thr Glu Arg Asn Leu Gln Asp Ala Gln Lys Phe Ile Leu Met His Glu
370 375 380
Val Val Gln Pro Val Thr Thr Val Pro Ser Phe Met Glu
385 390 395

<210> 15
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC15398

<400> 15
gtggctcagg tttgtaag 18

<210> 16
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC15397

<400> 16
gtggctcagg tttgtaag 18

<210> 17
<211> 963
<212> DNA
<213> Artificial Sequence

<220>
<221> variation
<222> (1)...(963)
<223> Each N is independently any nucleotide.

<223> Degenerate nucleotide sequence encoding the ZSMF-3 polypeptide of SEQ ID NO:7.

<400> 17

atgmgnccng cngcnytnyt nytnntgyytn acnytnytnc aytgygcngg ngcnggntty	60
ccngargayw sngarccnat hwsnathwsn cayggnaayt ayacnaarca rtayccngtn	120
ttygtngnc ayaarcnngg nmgnayt acnacrmgn aymgnytna yathcaratg	180
athatgatha tgaaymgnac nytnaygt gcngcnmgng aycayathta yacngtngay	240
athgayacnw sncayacnra rgarathtay tgywsnaara arytnacntg gaarwsnmgn	300
cargcngayg tngayacntg ymgnatgaar ggnaarcaya argaygartg ycayaayt	360
athaargtny tnytnaaraa raaygagay acnytnytna tntgyggnc naaygcntt	420
aayccnwsnt gymgnayta ymgnatnayt acnytnytna cnytggngt ygarttysn	480
gnatggcnm gntgycnta ygaygcnaar caygcnaaya thgcnytna tgcngayggn	540

aarytntayw sngcnacngt nacngaytty ytngcnathg aygcngtnat htaymgnwsn 600
ytnggngayw snccnacnyt nmgnacngtn aarcaygayw snaartggyt naargarccn 660
tayttygtnc argcngtna ytagggngay tayathtayt tyttypyng ngarathgcn 720
gtngartaya ayacnatggg naargtngtn tycnmgng tngcncargt ntgyaaraay 780
gayatgggng gnwsncarmg ntgytngar aarcartgga cnwsntyyt naargcnmgn 840
ytnaaytgyw sngtncnng ngaywsncay ttytaytta ayathytnca rgcngtnacn 900
gaygttnathm gnathaaygg nmgnaygtn gtnytnacn cnytwsnac nccntayaay 960
wsn 963

<210> 18

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC15523

<400> 18

aagtcccccc gctgaaccta ct

22

<210> 19

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16443

<400> 19

atcttctggg aaaccagccc cagcaaagtg tagcagtgtg aaatatacgca

50

<210> 20

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16462

<400> 20

gtgttagcgt gtgaaatata gcagcaaggc ttctgaccc atagtagtgg cgatccaggg
aattccaaca c

60

71

<210> 21

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14454

<400> 21

tcgtccaaacg actataaaga gggcaggctg tcctctaagc gtcacc

46

<210> 22

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14392

<400> 22

ggctgtcctc taagcgtcac cccggatcg ccactgtgtt ggaattc

47

<210> 23

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14393

<400> 23

atgccacccc agatcctcta gattattcca tggcatgta ttcttc

46

<210> 24

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14395

<400> 24

ggcactgggg aggggtcaca gggatgccac cccagatcct ctaga

45

<210> 25

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16441

<400> 25

gggcgtgatg ttgcctggc aacgtttct acaccttata acaggaaaga atacatgcc 60
atgaaataat c 71

<210> 26

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16444

<400> 26

gcagttacag atgtgattcg tatcaacggg cgtgatgttg tcctggca 48

<210> 27

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14397

<400> 27

ccgagttcca acgcggccgt agagaggagt atatgcctat ggag 44

<210> 28

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14396

<400> 28

gtttcgccca gccaggaaat ccatgccgac tcccaacgcg gccgt 45

<210> 29

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16463

<400> 29

gccatgcgaa atactgattg gctcagaatc ttctggaaa ccggatccct ccataggcat 60
atactccctct c 71

<210> 30

<211> 53

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16442

<400> 30

cacaaacacc ggatactgtt ttgtatagtt gccatgcgaa atactgattg gct 53

<210> 31

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14455

<400> 31

cacgagaggc actggggagg ggtcacaggg atgccacccc agatcc 46

<210> 32

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC14394

<400> 32

ggcactggag tggcaacttc cagggccagg agaggcactg gggagg 46

<210> 33

<211> 75

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16461

<400> 33

gggcgtgatg ttgtcctggc aacgtttct acaccttata acctggtggc atccggatct
gggaggtAAC tcgag

60

75

<210> 34

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC447

<400> 34

taacaatttc acacagg

17

<210> 35

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16296

<400> 35

cctgtctaga tttccatgtc

20

<210> 36

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16292

<400> 36

gacacatcac acacggaaga

20

<210> 37

<211> 381

<212> PRT

<213> Mus musculus

<400> 37

Ala Ala Ala Ala Cys Pro Ser Met His Arg Lys Gly Ser Arg Glu Ala
 1 5 10 15
 Ser Pro Pro Ala Glu Pro Thr Met Arg Pro Ala Ala Leu Leu Leu Cys
 20 25 30
 Leu Thr Leu Leu His Cys Ala Gly Ala Gly Phe Pro Glu Asp Ser Glu
 35 40 45
 Pro Ile Ser Ile Ser His Gly Asn Tyr Thr Lys Gln Tyr Pro Val Phe
 50 55 60
 Val Gly His Lys Pro Gly Arg Asn Thr Thr Gln Arg His Arg Leu Asp
 65 70 75 80
 Ile Gln Met Ile Met Asn Arg Thr Leu Tyr Val Ala Ala Arg
 85 90 95
 Asp His Ile Tyr Thr Val Asp Ile Asp Thr Ser His Thr Glu Glu Ile
 100 105 110
 Tyr Cys Ser Lys Lys Leu Thr Trp Lys Ser Arg Gln Ala Asp Val Asp
 115 120 125
 Thr Cys Arg Met Lys Gly Lys His Lys Asp Glu Cys His Asn Phe Ile
 130 135 140
 Lys Val Leu Leu Lys Lys Asn Asp Asp Thr Leu Phe Val Cys Gly Thr
 145 150 155 160
 Asn Ala Phe Asn Pro Ser Cys Arg Asn Tyr Arg Val Asp Thr Leu Glu
 165 170 175
 Thr Phe Gly Asp Glu Phe Ser Gly Met Ala Arg Cys Pro Tyr Asp Ala
 180 185 190
 Lys His Ala Asn Ile Ala Leu Phe Ala Asp Gly Lys Leu Tyr Ser Ala
 195 200 205
 Thr Val Thr Asp Phe Leu Ala Ile Asp Ala Val Ile Tyr Arg Ser Leu
 210 215 220
 Gly Asp Ser Pro Thr Leu Arg Thr Val Lys His Asp Ser Lys Trp Leu
 225 230 235 240
 Lys Glu Pro Tyr Phe Val Gln Ala Val Asp Tyr Gly Asp Tyr Ile Tyr
 245 250 255
 Phe Phe Phe Arg Glu Ile Ala Val Glu Tyr Asn Thr Met Gly Lys Val
 260 265 270
 Val Phe Pro Arg Val Ala Gln Val Cys Lys Asn Asp Met Gly Gly Ser
 275 280 285
 Gln Arg Val Leu Glu Lys Gln Trp Thr Ser Phe Leu Lys Ala Arg Leu
 290 295 300
 Asn Cys Ser Val Pro Gly Asp Ser His Phe Tyr Phe Asn Ile Leu Gln
 305 310 315 320
 Ala Val Thr Asp Val Ile Arg Ile Asn Gly Arg Asp Val Val Leu Ala
 325 330 335

Thr Phe Ser Thr Pro Tyr Asn Ser Ile Pro Gly Ser Ala Val Cys Ala
340 345 350
Tyr Asp Met Leu Asp Ile Ala Asn Val Phe Thr Gly Arg Phe Lys Glu
355 360 365
Gln Lys Ser Pro Asp Ser Thr Trp Thr Pro Val Pro Asp
370 375 380

<210> 38

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC17960

<400> 38

gtatacggcc ggccaccatg aggtcagaag cc

32

<210> 39

<211> 32

<212> DNA

<213> Artificial Sequence

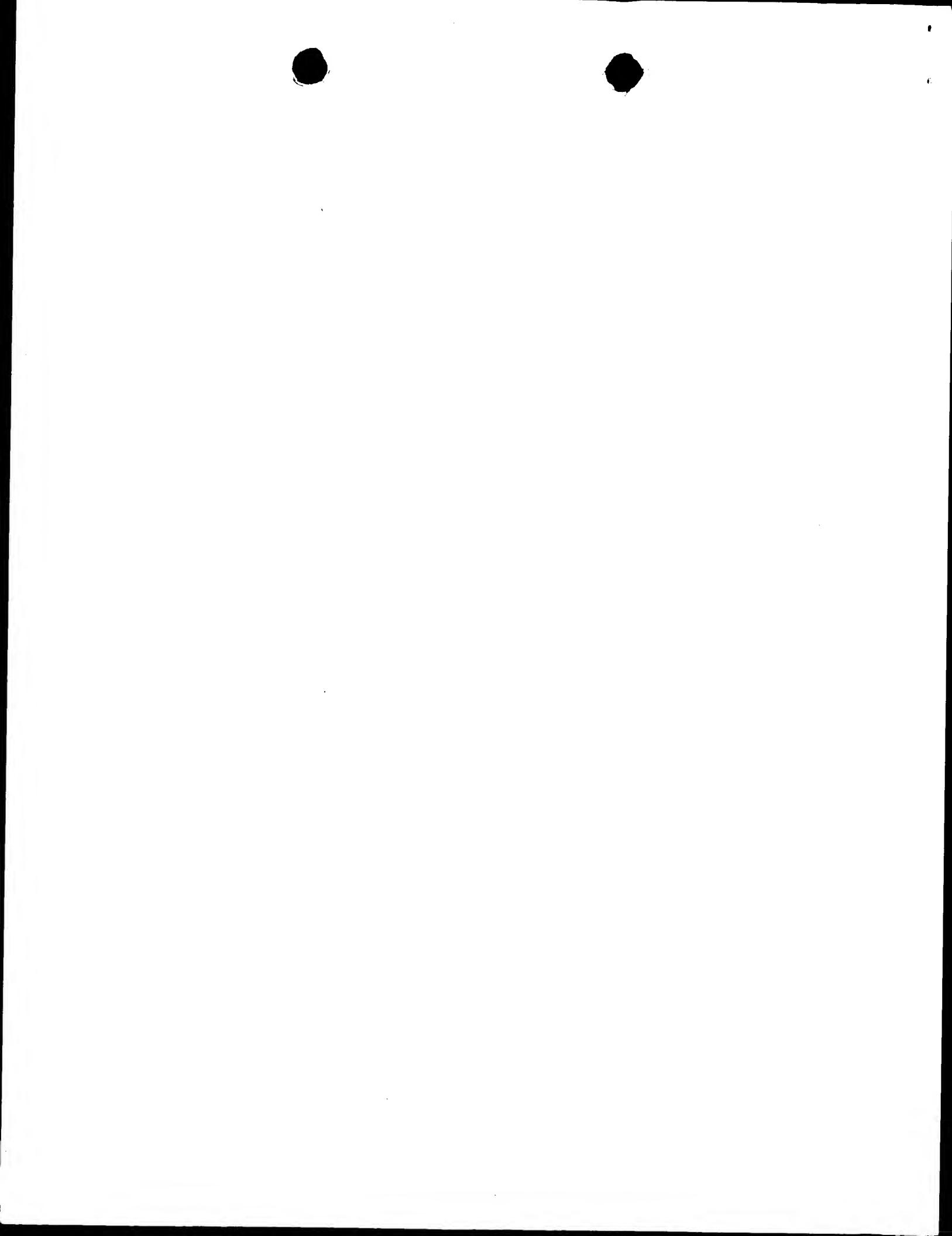
<220>

<223> Oligonucleotide ZC17961

<400> 39

cgcgcggcgc cgccttacct gttataaggt gt

32



PCT

UNITED INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, 15/62, C07K 14/47, 16/18, A61K 38/17, C12N 5/10		A3	(11) International Publication Number: WO 99/32622 (43) International Publication Date: 1 July 1999 (01.07.99)
(21) International Application Number: PCT/US98/26793		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 16 December 1998 (16.12.98)		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data: 08/994,494 19 December 1997 (19.12.97) US		(88) Date of publication of the international search report: 10 September 1999 (10.09.99)	
(71) Applicant: ZYMOGENETICS, INC. [US/US]; 1201 Eastlake Avenue East, Seattle, WA 98102 (US).			
(72) Inventor: HOLLOWAY, James, L.; 835 N.E. 89th Street, Seattle, WA 98115 (US).			
(74) Agent: LINGENFELTER, Susan, E.; ZymoGenetics, Inc., 1201 Eastlake Avenue East, Seattle, WA 98102 (US).			
(54) Title: SEMAPHORIN ZSMF-3			
(57) Abstract <p>Semaphorin polypeptides, polynucleotides encoding the polypeptides, and related compositions and methods are disclosed. The polypeptides include a soluble, truncated semaphorin that is highly expressed in stomach tissue. The polypeptides may be used within methods for detecting receptors that mediate neurite outgrowth, and may also have immunological value.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/US 98/26793

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/62 C07K14/47 C07K16/18 A61K38/17
C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>ZHOU L ET AL: "Cloning and expression of a novel murine semaphorin with structural similarity to insect semaphorin I" MOLECULAR AND CELLULAR NEUROSCIENCES., vol. 9, no. 1, January 1997 (1997-01), pages 26-41, XP002109255 SAN DIEGO, US ISSN: 1044-7431 cited in the application abstract figure 1 95.9% identity in 320 AA overlap with SEQ ID 2 97.3% identity with the fragment of SEQ ID 2 spanning from residue 20 to residue 321 ---</p> <p style="text-align: center;">-/-</p>	1-8, 13-15

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

15 July 1999

30/07/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Lejeune, R

INTERNATIONAL SEARCH REPORT

National Application No

PCT/US 98/26793

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	WO 98 27205 A (GENETICS INST) 25 June 1998 (1998-06-25) 100% identity in 320 AA overlap between SEQ ID 8 of WO9827205 and SEQ ID 2 of the application ----	1-8, 13-15
A	KOLODKIN ET AL: "Semaphorins: Mediators of repulsive growth cone guidance" TRENDS IN CELL BIOLOGY, vol. 6, no. 1, 1 January 1996 (1996-01-01), pages 15-22, XP002084499 ISSN: 0962-8924 cited in the application -----	1-30

INTERNATIONAL SEARCH REPORT

Information on patent family members

In / Application No
PCT/US 98/26793

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9827205 A	25-06-1998	AU 5706198 A	15-07-1998

